STUDIES IN LANGUAGE COMPANION SERIES 195

Changing Structures

Studies in constructions and complementation

Edited by
Mark Kaunisto
Mikko Höglund
Paul Rickman

JOHN BENJAMINS PUBLISHING COMPANY

Changing Structures

Studies in Language Companion Series (SLCS)
ISSN 0165-7763

This series has been established as a companion series to the periodical *Studies in Language*.

For an overview of all books published in this series, please see *http://benjamins.com/catalog/slcs*

Volume 195

Changing Structures. Studies in constructions and complementation
Edited by Mark Kaunisto, Mikko Höglund and Paul Rickman

Changing Structures

Studies in constructions and complementation

Edited by

Mark Kaunisto
University of Tampere

Mikko Höglund
Stockholm University

Paul Rickman
University of Tampere

John Benjamins Publishing Company
Amsterdam / Philadelphia

DOI 10.1075/slcs.195

Cataloging-in-Publication Data available from Library of Congress.

ISBN 978 90 272 0054 9 (HB)
ISBN 978 90 272 6422 0 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

Part 3. The emergence of new patterns

Acknowledgments

This volume is the product of a symposium that was held at the University of Tampere in October 2015. The theme of that one-day event was complementation, and it was held in honour of Professor Juhani Rudanko, who, at the time, was retiring from his post at the University. Half of the contributors to our volume share the good fortune of having been Juhani's students at some point in his career; the other half have all had the pleasure of being associated with him in other ways, as a result of a common interest in complementation, or one of the other areas he is involved in. Thus the present volume, a collection of papers presented at the symposium, along with some invited contributions, is dedicated to Juhani, a gentleman and a scholar, in recognition of his substantial contributions to the field of complementation.

Thanks must go to the English Degree Programme at the University of Tampere, for the financial support which made the symposium possible. We also thank Juhani Klemola, Päivi Pahta, and Peter Slomanson, who, along with two of the editors, Mark Kaunisto and Paul Rickman, served on the organising committee. Mikko Höglund, Juhani's former student and colleague, and one of the speakers at the symposium, kindly agreed to step in as a third editor of this volume. Emilia Luukko took care of all secretarial duties for the symposium with great efficiency and goodwill.

A number of people were called upon to act as external referees for the papers in this collection, and to these people, who allowed us to impose upon their time and make use of their expertise, we extend our warmest thanks: Laurie Bauer, Cristiano Broccias, Mark Davies, Hendrik De Smet, Patrick Duffley, Thomas Egan, Teresa Fanego, Olga Fischer, Tuomas Huumo, Urpo Nikanne, Gerold Schneider, Ian Smith, An Van linden, and Bao Zhiming. Naturally, the editors and the authors are responsible for all final decisions regarding the content of the chapters.

We are indebted to John Benjamins for including the volume as part of the Studies in Language Companion series. It has been a pleasure working with the people at Benjamins, with their cordial and uncomplicated approach to business matters. We thank Kees Vaes in particular for all the assistance.

The preparation of a volume like this is obviously the concerted effort of a large group of people, and those who provide the papers themselves are at the heart of the process. It has been a pleasure working with all of our contributors; it is mainly thanks to them, their professionalism, and the high quality of their work even in the early stages, that we have been able to keep roughly to our original production schedule.

Structures, patterns, constructions – studying variation and change in lexico-grammar

Mark Kaunisto
University of Tampere

As is typical of language in general, in the study of the structure of language many terms are often used for one and the same thing. This also applies to the very concept of *structure*. Combinations of words or phrases with other elements, constituting functional units, are occasionally referred to by linguists as "structures", "constructions", "patterns" – or, indeed, "combinations" or "units". The combinations may make up phrases, clauses or entire sentences, the structures of which may be considered in their own right. For example, the co-occurrences of adjectives and nouns within the same phrase may be of great interest, as certain types of members of these word classes are more likely to co-occur with each other than with other types.

There may be different reasons for the variety of expressions used for similar things in the linguistic literature, and in some cases authors may have clear preferences for using a particular term for particular types of elements. For one thing, some theoretical approaches have such terms in their names, e.g. Pattern Grammar (see e.g. Hunston & Francis 1999) and Construction Grammar (see e.g. Goldberg 1995 or Hilpert 2014), which have their own sets of theoretical positions as well as practical tools employed in the description of language, and the terms used are reflective of how different linguistic elements and their roles are understood or defined. There are instances in which the introduction of new terms for relatively closely related concepts can be regarded as justified. For example, while the term *collocation* is generally used of the adjacency or co-occurrence of lexical items – and the study of collocations with corpus linguistic methods has undoubtedly changed the face of language study in several applied fields, such as lexicography – one can extend the analysis of the co-occurrence of lexical items with each other to that of lexical items with grammatical structures with more or less clearly definable functions. Examples in this regard would be ditransitive structures (i.e. patterns of the type of *He gave me a book*) or verbs with *to*-infinitival clausal complements (*He remembered to mention it*). For the observation of groups of lexical items that are more (or less) likely to occur in

DOI 10.1075/slcs.195.01kau

such structures, the term *collostruction* has also been introduced (e.g. Stefanowitsch & Gries 2003; Hilpert 2014).

Although the use of expressions such as "patterns", "structures" and "constructions" may reflect adherence to a particular theoretical approach in linguistics, it is still possible that some authors use these terms in free variation, perhaps for the simple reason of avoiding repetition. Furthermore, even greater differences in the use of terminology in general should not override the overall goal of describing the nature of language, which is shared by all approaches. It is worth noting, as observed by McEnery and Hardie (2012: 162) in their discussion on the differences between so-called neo-Firthian linguistics and functional linguistics, that the "different camps often arrive at similar conclusions as a result of their studies."

The observation made by McEnery and Hardie also serves as one of the guiding principles in the compilation of contributions to the present volume, which brings together 11 chapters by authors in various stages of their academic careers examining various aspects in the occurrence and variation of different structural elements in English and other languages. The aim has not been to focus strictly on any particular theoretical framework; instead, different frameworks are applied, referred to and commented on. Likewise, the methods employed in these studies vary, with some of them basing their argumentation on corpus linguistic evidence, while others discuss details of theoretical issues in a contemplative fashion. However, the collection is not a random mixture of loosely connected studies; there are clear themes that tie the chapters together. These include:

i. the description of the structure and functions of different types of complementation patterns
ii. the examination and/or compilation of diachronic and synchronic corpus data in the study of complementation patterns, and
iii. the study of variation of usage in different registers and regional varieties.

These themes will be discussed next in greater detail.

i. The study of complementation patterns

In the study of the English language, a great deal of attention has been paid to the complementation patterns of verbs, adjectives, and nouns. These items can vary as regards their selection of clausal and/or non-clausal complements, with some items selecting, for example, non-finite as well as finite clausal complements. In the first half of the twentieth century, grammarians such as Kruisinga (1922), Poutsma (1929), Curme (1931), and Jespersen (1940) made observations about the behaviour of verbs, adjectives, and nouns in this regard, making great inroads into the description of the occurrence and use of various patterns.

The continued efforts in this field have benefited from the introduction of sizable electronic corpora in ways that cannot be overstated – corpora have made it possible to examine both large-scale diachronic developments as well as analyse the significance of individual factors in the selection of complement types. As regards historical trends, certain key changes in the wider system of complementation have been perceived by several scholars as having occurred over the last couple of centuries, a set of developments to which Rohdenburg (2006) applied the term The Great Complement Shift. This includes the spread in the use of gerundial non-finite complement clauses at the expense of infinitival complements (see e.g. Fanego 1996, 1997; Vosberg 2003, Egan 2008, and Rudanko 2015, 2017).

In addition to examining which kinds of complements are attested in connection with different matrix verbs, adjectives and nouns, attention has also been paid to different types of factors that may influence the choice between alternative complementation patterns. Such factors may be syntactic or semantic as regards their nature. For example, in cases where matrix predicates may select different types of complements, features contributing to the complexity of the syntactic structure of the sentence – such as passivization, or extraction of a syntactic element – are likely to result in the use of grammatically more explicit complementation patterns, e.g. infinitives instead of gerundial complements. This tendency has been termed the Complexity Principle (Rohdenburg 1996; see also Vosberg 2006; Mondorf 2009; and Berlage 2014). As regards semantic factors, Rudanko (2010, 2011, 2017) has examined the role of agentivity in connection with the selection of complements, a tendency which is referred to as the Choice Principle. The influence of these factors are also examined in the present volume in the contributions by Rohdenburg and Höglund.

The studies in the present volume focus mainly on the complementation of verbs, and include the following types of patterns:

- matrix verb + finite/non-finite clausal complement where the subject of the matrix verb is also the subject of the lower-level verb (e.g. the English *propose* + *that*-clause/*to*-infinitive/*-ing* clause, the Finnish *rakastaa* ('love') + infinitive)
- matrix verb + NP + non-finite clausal complement, with the object NP being the subject of the verb in the complement clause (*let/allow* + NP + *to*-infinitive)
- matrix verb (+ NP) + preposition + non-finite clausal complement (e.g. the intransitive particle verb *hold off* + (*from*) *-ing*, *put* NP *off* (*from*) *-ing*, *talk/convince/force*/etc. + NP + *into* + *-ing*)
- matrix verb (+ preposition) + NP (i.e. verbs taking either a direct object or a prepositional object)

In addition to verbs, adjectives and nouns can take clausal complements; these also receive attention in the present volume (variation between *to*-infinitive and *of -ing* complements with the adjective *ashamed,* and the omissibility of *that*-complementizer with nouns).

ii. The use of corpora

The availability of large collections of authentic data has changed the face of the study of complementation, and the wide variety of corpora we now have access to allows a greater range of possibilities for the examination of different types of variation. For the present volume, some of the contributors make good use of a number of older, well-known corpora, and others use some of the newer corpora that have recently been compiled specifically for the purpose of studying individual regional varieties of English. To give an idea of the breadth and sheer volume of language data investigated, it may be helpful to list all the corpora and text archives examined in the papers included in the present volume:

- Nineteenth Century Fiction database (37.5 million words, British English)
- The Corpus of Late Modern English Texts (CLMET, version 1; 1710–1920, totaling approximately 10 million words, British English)
- The Corpus of Historical American English (COHA; 400 million words, 1810s–2000s)
- The Lancaster-Oslo/Bergen corpus (LOB; 1 million words, 1961 British English)
- The British National Corpus (BNC; 100 million words, 1960–1994)
- The Corpus of Contemporary American English (COCA; 520 million words, 1990–2015)
- The Corpus of New Zealand Newspaper English (CNZNE, 1995–98 and 2010–12, approximately 100 million words)
- The Global Web-based Corpus of English (GloWbE, GB and US sections; approximately 774 million words)
- The NOW Corpus (News on the Web; 3.8 billion words of newspaper data; 20 regional varieties of English)
- CD-ROM issues of British and American newspapers (sets of yearly issues between 1990 and 2005; altogether approximately 3.1 billion words)
- newspaper data compiled from the *Straits Times* newspaper (from the years 1951, 1961, 1971, 1981, 1991, 2001 and 2011; 158 million words; Singapore English)
- The Archives of Finnish Syntax (Finnish newspapers and magazines from the 1990s and 2000s; 80 million words)
- Suomi24 Corpus (database of Finnish online discussions; 2.6 billion words)

As can be seen in the list of corpora above, they vary considerably as regards their size as well as the registers and varieties that they represent. Ultimately the choice of analyzing a particular corpus depends largely on the frequency of the examined subject at hand. Even corpora such as the one-million-word LOB corpus, compiled at the dawn of the modern corpus era, are still highly relevant in the exploration and description of certain areas of language use, as the study by Rohdenburg in the present volume attests.

Considering the methodological challenges in analyzing corpus data, the authors have taken great care to clarify what kinds of conclusions can be justifiably drawn based on the findings, as due caution is often needed. Both Callies and Hoffmann, for example, observe some of the potential problems in the contents of the GloWbE corpus, reminding the reader that some sections of the corpus may not entirely reflect usages of the particular varieties. In similar fashion, authors do well in pointing out the idiosyncratic uses affecting the frequencies in their results, as well as giving examples of some false positives among the search hits. Such reminders are of great importance, as knowledge of the contents of the corpora and close scrutiny of the concordance lines form the solid basis on which the analysis of the data can be built. To further ensure the reproducibility of the studies, the authors also detail the actual search queries or query strings. When relevant, the significance of the findings has been assessed with statistical methods.

iii. Examining regional variation in the use of complementation patterns

In recent years, the description of grammatical patterns found in different regional varieties of English has taken great steps forward. The first electronic corpora represented the major varieties, British and American English, but with new, specialized corpora other varieties and genres can be examined in more detail. The study of World Englishes with representative corpus data makes it possible to contribute to the theoretical models used for the classification and evolution of varieties, such as Kachru's (1985) three-circle model with the Inner, Outer, and Expanding Circle varieties, and Schneider's (2003, 2007) five-phase model outlining the steps in the evolution of a new regional variety. The present volume contains studies on the differences between the main varieties (the contributions by Rohdenburg and Callies), as well as ones focusing on a single regional variety, with comparisons made to more dominant varieties (Rickman & Kaunisto on New Zealand English, and Hoffmann on Singapore English).

The studies on individual varieties provide us with greater insights into the trend lines in the developments of colonial varieties and/or varieties of English spoken as a second or foreign language. With regard to the changes seen between British English and other varieties, differences have been also examined in terms of retention of older forms (also referred to as "colonial lag") or innovation (or

"colonial lead") (see e.g. Görlach 1987; Montgomery 2001; Hundt 2009). As is noted in the studies in the present volume, characterizing the differences as reflecting one or the other type in relation to another variant is a complex issue. This also applies to the process of new, emerging complementation patterns or rival patterns in languages other than English; in the present volume, two studies focusing on Finnish (Hietaranta) and Sri Lankan Malay (Slomanson) discuss the possible factors behind the introduction of novel structures into the grammatical system of these languages.

The volume has been structured to present the studies according to three main themes as regards their main focus and/or methodological approaches, namely:

1. the semantic and functional descriptions of constructions
2. corpus studies describing the distribution of complementation patterns
3. studies on innovative patterns.

This division is, of course, not entirely clear-cut; instead, in many instances the discussion in the studies is relevant from the perspective of more than just one of the above-mentioned themes.

Part 1, 'Semantic descriptions of constructions' includes studies in which the emphasis is on the semantic content of the structures under investigation. The studies discuss and comment on the ways in which relationships between linguistic structures and meaning are described, commenting on the theoretical frameworks of Construction Grammar and Cognitive Grammar. **Patrick Duffley** discusses the causative constructions *talk* NP *into -ing* and *convince* NP *to* + infinitive, and offers critical viewpoints on some of the earlier accounts of the structures, and in particular on the Construction Grammar approach. Duffley proposes that the structures themselves on the whole do not form unique combinations to which specific meanings are assigned but that the verbs found in such structures are predictable based on the meanings of the components of the structure. Thus in the case of the V NP *into -ing* pattern, it is the fundamental meaning of the preposition *into* that explains much of the attested usage of the pattern. **Jouni Rostila** also makes observations on the Construction Grammar approach into the analysis of structures in his contribution on the argument structure construction termed the Rely On construction. Commenting on Goldberg's (2014) account of the structure, Rostila finds Goldberg's model semantically too broad and vague, a criticism seen also in previous reactions to argument structure constructions (or ASCs) in the literature. Rather than rejecting the idea of ASCs altogether, he offers his own model to describe the construction, influenced by contrastive observations of similar structures in German. In his conclusion, Rostila generally recommends the contrastive method as a possible path to making inspired observations and analyses

of constructions, as long as principles relevant to the target languages are kept in mind. **Thomas Egan**'s paper examines the verb *let* and why it is very seldom found in the passive permissive construction as in *They should not be let (to) leave so soon.* Egan compares this characteristic of *let* with other verbs, such as *allow* and *permit*, which are commonly found in comparable passive constructions. Using data from the British National Corpus, the Corpus of Contemporary American English, and the Corpus of Historical American English to back up his observations, Egan discusses the differences in the force dynamic relations of the verbs as well as in the fundamental meanings between *to* infinitives and bare infinitives. In his discussion of the meanings of verbs expressing permission, Egan points out that it is used to signal different types of permission, namely 'barrier-removal' and 'non-imposition of barrier', and their negative counterparts 'barrier-retention' and 'imposition of barrier'. He notes that differences between *let* and *allow* can be seen in this regard even in their usage as matrix verbs both in the active and passive. With the semantic feature encoded typically by *let* in the active being 'non-imposition', its passive form counterpart would be largely uninformative, which thus explains why it is not often used.

Part 2 of the volume, 'Variation and change in complementation patterns' presents corpus-based studies observing issues relating to complementation from a number of different perspectives. **Paul Rickman and Mark Kaunisto** examine the host of verbs used in the V NP *into -ing* pattern in New Zealand English, comparing the type and token frequencies as well as semantic characteristics of such verbs against similar studies conducted on American English. They examine New Zealand newspaper data from two periods, 1995–98 and 2010–12 and observe whether there are perceivable changes in the uses of the pattern. The results suggest that the productivity of the pattern has continued to be relatively high, as is also the case in present-day American English. Rickman and Kaunisto also highlight matrix verbs found in the data which were not previously attested in earlier surveys and which thus may be considered to be localized, New Zealand-based innovations in the transitive *into -ing* pattern. **Mikko Höglund** focuses on the different complements selected by the matrix adjective *ashamed*, and as data, he uses the Corpus of Historical American English and the Corpus of Contemporary American English. Höglund conducts a diachronic survey into the frequencies of different types of complements, including *to*-infinitives, *that*-clauses, *of -ing* patterns, *of* + NP patterns, as well as instances without complements. Höglund finds that, in accordance with the general tendencies predicted by the Great Complement Shift, the frequency of the *to*-infinitive complements with *ashamed* has been declining, but this does not seem to have coincided with an increase in the use of the gerundial *of -ing* pattern. In addition, Höglund investigates different factors that might have an effect on the selection between two non-finite complement

types, the *to*-infinitive and *of -ing*, and observes that, for example, copula or negation in the complement clause attract the gerundial complement type. In a paper based on her MA thesis, **Veera Saarimäki** examines the sentential complementation patterns of the matrix verb *propose* from the late 18th century to the 1990s, using the Corpus of Late Modern English Texts and the British National Corpus as sources of data. She first presents an overview of how the use of the verb with different types of complements has been described in dictionaries and grammars, and then looks at diachronic corpus data to track changes in the proportions of different sentential complements with the verb. In terms of the use of different types of complements, the diachronic changes in connection with the verb *propose* are not in line with the Great Complement Shift as *to*-infinitival complements appear to have increased their proportional frequency while *-ing* clauses have become less frequent. In addition to the occurrences of different complementation patterns, Saarimäki also makes observations on the use of *propose* in relation to the notion of control, i.e. the notion that the lower-level verbs have understood subjects which are co-referent with (mostly) either the subject or object of the higher-level clause. Saarimäki notes that while the majority of occurrences of the verb with clausal complements manifest subject control, some examples of object control, PP object control, as well as unspecified control, can be detected as well. **Günter Rohdenburg** contributes to the volume with a study on the factors influencing the presence of a set of optional complement markers, including the complementizer *that*, the infinitival marker *to*, the modal verb *should* after mandative predicates, and the preposition *from* when introducing gerundial complements after particle verbs. In addition, he looks into the optionality of prepositions when they introduce interrogative complement clauses. With each type of complement marker, Rohdenburg examines the influence of factors adding to the grammatical complexity of the structures in which the complement markers are found. In doing so, Rohdenburg makes use of a number of corpora and databases of British and American English. For the most part, the focus is on the influence of passivization on the omissibility of the complement markers, but other factors are also considered. In the conclusion, Rohdenburg discusses what the findings contribute to the explanatory power of the Complexity Principle in the selection between alternative complementation patterns and forms. Concluding Part 2, **Marcus Callies** examines a tendency seen in major varieties of English in connection with a number of verbs increasingly selecting direct object noun phrases instead of prepositional objects, as in *They fleed [from] the extreme weather conditions* or *People were protesting [against] the idea*. Callies observes that earlier commentaries and studies on this issue offer contradictory views on the historical stages of the tendency and on whether it is nowadays more common to omit the preposition in

such cases in British or American English. In an earlier study, Rohdenburg (2009) had examined this tendency in connection with a group of antagonistic verbs (e.g. *fight, protest,* and *offend*) and verbs of leaving (e.g. *flee, escape,* and *depart*). Callies studies three other verbs – *graduate, impact,* and *shop* – that have shown signs of allowing omission of the preposition (or "direct transitivization"), and focuses on the variation between the direct object and prepositional object patterns in British and American English corpus data.

The papers in Part 3 of the volume, 'The emergence of new patterns', focus on the processes of innovation in the use of patterns of complementation. While much of the study of English complementation patterns deals with the occurrence of the patterns in main varieties of English, increasing attention is also given to variation seen in non-native Englishes, used as a second or a foreign language. Furthermore, structural analyses of languages other than English can be conducted by observing parallels between English and the other languages – even non-Indo-European ones – a method which has only become more relevant from the point of view of language contacts in a world where the influence of English is more and more pervasive. Like Callies, **Sebastian Hoffmann** also looks into the question of particle verb usage, focusing on the ESL variety of Singapore English. This variety manifests a trend that is in a sense opposite to that examined by Callies in that Singapore English data contains instances of prepositional verbs which in main varieties would take a direct object without a preposition (e.g. *discuss (about) something* and *emphasise (on) something*). In order to examine the diachronic developments in the use of such items, Hoffmann compiled a 158-million-word corpus containing data from one Singaporean newspaper, the *Straits Times,* from the period spanning the years 1951 and 2011. The closer studies of four prepositional verbs attested in Singapore English (*enter into, await for, request for*, and *leverage on*), also examined in historical and present-day native English databases and corpora, show that different reasons may be suggested for why prepositions are used in connection with these verbs in SingE. The evidence of the use of the preposition as instances of retention of older, nowadays obsolete forms of British English is not very prominent. Instead, analogy and innovation in an ESL context appear to be more likely explanations, and Hoffmann points out that there is room for more work on the phenomenon on the whole.

While new, more specialized corpora are being compiled at an ever-increasing speed, it is worth noting that there are still many languages for which representative corpora are not available. The paper by **Peter Slomanson** shows how the characterization and description of structural changes seen in rare languages benefits from observations of parallel developments across different languages. Basing his observations on his ongoing fieldwork involving interviews and recordings made

in rural Sri Lanka, Slomanson discusses the development of infinitival complementation in Sri Lankan Malay, a South Asian contact language which earlier did not show finiteness contrast, and points to similarities that can be perceived in how comparable structures emerged in Old English. The root in the rise of the infinitive in Sri Lankan Malay is seen to have emerged from a prepositional phrase with a dative marker. Contrasts and contacts between languages as well as sociolinguistic factors are also examined by **Pertti Hietaranta**, who discusses a syntactic change occurring in Finnish today, which allows the verb *rakastaa* ('to love') as a matrix verb to be followed by an infinitival verb form, as in *rakastan puhua* ('I love to speak') instead of the standard structure *rakastan puhumista* ('I love speaking'). He observes that although the influence of English into Finnish through translation is one possible explanation for the growing popularity of the new construction, ideological factors, new means of communication and fashion may, in fact, play a more prominent role.

One final point which deserves to be made has to do with how the studies in the present volume vary as regards the scope of examination. There are both surveys which have the aim of detecting broader patterns of use that are common to larger groups of matrix words (e.g. the paper by Rickman and Kaunisto on verbs selecting NP + *into -ing* complements), and studies with a focus on a single matrix word. Studies of both such types make valid contributions to the overall knowledge on complementation, although it is important to note the comment made by Leech et al. (2009) on studies with a broad or narrow focus:

> If we decide to focus on a specific non-finite complement structure – such as, say, the *to*-infinitive clause or the gerund with possessive/genitive modifier – we will find these structures serving a large variety of functions, with most of them not being involved in current diachronic change. If, on the other hand, we decide to focus on more specific constructions – combinations of particular superordinate predicates and particular patterns of complementation (such as, for example, variation between infinitives and gerunds with *accustomed to*) – we can easily home in on areas of ongoing diachronic change, without, however, being able to correlate individual shifts in usage preferences with general trends in the evolution of the system of English non-finite verbal forms. (Leech et al. 2009: 181)

In other words, there are pros and cons inherent in both approaches. However, individual matrix predicates may show tendencies of use which can go against those of similar words, and explanations may be found in the semantic component of the words themselves. In-depth analyses of individual items may therefore reveal important aspects affecting the linguistic behaviour of the items that may be unique to those very items. This knowledge can then be taken into account when making generalizations on a broader scale.

References

Berlage, Eva. 2014. *Noun Phrase Complexity in English*. Cambridge: CUP. doi:10.1017/CBO9781139057684

Curme, George O. 1931. *Syntax*. Boston MA: Heath.

Egan, Thomas. 2008. Emotion verbs with to-infinitive complements: From specific to general predication. In *Selected Papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006, Vol. I: Syntax and Morphology* [Current Issues in Linguistic Theory 296], Richard Dury, Maurizio Gotti & Marina Dossena (eds), 223–240. Amsterdam: John Benjamins.

Fanego, Teresa. 1996. On the historical development of English retrospective verbs. *Neuphilologische Mitteilungen* 97: 71–79.

Fanego, Teresa. 1997. On patterns of complementation with verbs of effort. *English Studies* 78(1): 60–67. doi:10.1080/00138389708599061

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Goldberg, Adele E. 2014. Fitting a slim dime between the verb template and argument structure construction approaches. *Theoretical Linguistics* 40(1–2): 113–135.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8(1): 41–60. doi:10.1075/eww.8.1.05gor

Hilpert, Martin. 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2(2): 243–256. doi:10.1515/CLLT.2006.012

Hilpert, Martin. 2014. Collostructional analysis: Measuring associations between constructions and lexical elements. In *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy* [Human Cognitive Processing 43], Dylan Glynn & Justyna A. Robinson (eds), 391–404. Amsterdam: John Benjamins. doi:10.1075/hcp.43.15hil

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 13–37. Cambridge: CUP.

Hunston, Susan & Francis, Gill. 1999. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English* [Studies in Corpus Linguistics 4]. Amsterdam: John Benjamins.

Jespersen, Otto. 1940. *A Modern English Grammar on Historical Principles*, Part V: *Syntax*. Copenhagen: Ejnar Munksgaard.

Kachru, Braj B. 1985. Standards, codification, and sociolinguistic realism: The English language in the outer circle. In *English in the World: Teaching and Learning the Language and the Literature*, Randolph Quirk & H. G. Widdowson (eds), 11–30. Cambridge: CUP.

Kruisinga, Etsko. 1922. *A Handbook of Present-Day English*, Vols. 1–2, 3rd edn. Utrecht: Kemink & Zoon.

Leech, Geoffrey, Hundt, Marianne, Mair, Christian & Smith, Nick. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: CUP. doi:10.1017/CBO9780511642210

McEnery, Tony & Hardie, Andrew. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: CUP.

Mondorf, Britta. 2009. *More Support for More-support: The Role of Processing Constraints on the Choice between Synthetic and Analytic Comparative Forms* [Studies in Language Variation 4]. Amsterdam: John Benjamins. doi:10.1075/silv.4

Montgomery, Michael. 2001. British and Irish antecedents. In *The Cambridge History of the English Language*, *Vol. VI: English in North America*, John Algeo (ed.), 86–153. Cambridge: CUP. doi:10.1017/CHOL9780521264792.004

Poutsma, Hendrik. 1929. *A Grammar of Late Modern English*, 2nd edn. Groningen: Noordhoff.

Rohdenburg, Günter. 1996. The complexity principle as a factor determining grammatical variation and change in English. In *Language Use, Language Acquisition and Language History: (Mostly) Empirical Studies in Honour of Rüdiger Zimmermann*, Ingo Plag & Klaus Peter Schneider (eds), 25–44. Trier: Wissenschaftlicher Verlag.

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms: English from 1500–2000*, Christiane Dalton-Puffer, Nikolaus Ritt, Herbert Schendl & Dieter Kastovsky (eds), 143–166. Frankfurt: Peter Lang.

Rohdenburg, Günter. 2009. Nominal complements. In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 1–12. Cambridge: CUP.

Rudanko, Juhani. 2010. Explaining grammatical variation and change: A case study of complementation in American English over three decades. *Journal of English Linguistics* 38(1): 4–24. doi:10.1177/0075424209335495

Rudanko, Juhani. 2011. *Changes in Complementation in British and American English: Corpus-Based Studies on Non-Finite Complements in Recent English*. Houndmills: Palgrave Macmillan. doi:10.1057/9780230305199

Rudanko, Juhani. 2015. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Houndmills: Palgrave Macmillan.

Rudanko, Juhani. 2017. *Infinitives and Gerunds in Recent English: Studies of Non-Finite Complements with Data from Large Corpora*. London: Palgrave Macmillan.

Schneider, Edgar W. 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79(2): 233–281. doi:10.1353/lan.2003.0136

Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the World*. Cambridge: CUP. doi:10.1017/CBO9780511618901

Stefanowitsch, Anatol & Gries, Stefan T. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2), 209–243. doi:10.1075/ijcl.8.2.03ste

Vosberg, Uwe. 2003. The role of extractions and horror aequi in the evolution of *-ing*-complements in Modern English. In *Determinants of Grammatical Variation in English*, Günter Rohdenburg and Britta Mondorf (eds), 305–327. Berlin: Mouton de Gruyter. doi:10.1515/9783110900019.305

Vosberg, Uwe. 2006. *Die große Komplementverschiebung. Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.

PART 1

Semantic description of constructions

Talk into vs *convince to*

Talking as a cause leading to containment, convincing as a cause leading to a result

Patrick J. Duffley
Université Laval

This paper explores the causative constructions '*talk NP into* + *-ing*' vs '*convince NP to* + infinitive' by means of a collection of attested occurrences. It shows the connection between the characteristics described by Wierzbicka (1998), Gries & Stefanowitsch (2004) and Rudanko (2006) and the linguistically-signified semantic content involved in these structures. Wierzbicka's account and the related Construction Grammar approach are shown to be wanting on both the descriptive and explanatory levels due to a distancing from the level on which a stable relation exists between meaning and linguistic form. An explanation of the distribution and semantics of the two constructions is proposed based on Langacker's (1987) semiological principle, i.e. on the semantic content associated with each of the linguistic signs involved in these sequences.

Keywords: causatives; container metaphors; path metaphors; Construction Grammar; Natural Semantic Metalanguage; embodiment

1. Introduction

Causative constructions have generated an abundant body of literature over the past 40 years (cf. among others Aissen 1979; Bardzokas 2012; Baron 1974; Comrie & Polinsky 1993; Gilquin 2010; Givón 1975, 1980; Gries & Stefanowitsch 2004; Hollmann 2005; Kim & Davies 2015; Rudanko 2006, 2015; Shibatani 1976; Song 1996; Talmy 1976; Verhagen & Kemmer 1997; Wierzbicka 1998). The first task to be undertaken in the present study is to refine the observation of the empirical data using modern tools not available to earlier researchers such as Wierzbicka, in this case the Internet and on-line electronic corpora. Besides the World Wide Web, two corpora will be used for this study: the 553-million-word Bank of English (henceforth BOE) and the 450-million-word Corpus of Contemporary American English (COCA).

DOI 10.1075/slcs.195.02duf

Our second task will be to argue, *contra* a constructionist type of approach, that the observations made by the authors who have studied the causative '*talk NP into* + *-ing*' and '*convince NP to* + infinitive' sequences can all be explained by the meanings of the items of which these sequences are composed. This leads into the third more theoretical section of this paper, which will involve a discussion of what these findings imply for those who have attempted to apply a constructionist approach to the sequences examined here and of the limitations of this type of approach in general.

2. Refining the observation of the empirical data

The first major discussion of the semantics of the English causative constructions investigated in the present study concerns the structures '*get* + object + *to* + infinitive,' as exemplified by *I got Joe to unlock the door*, versus 'verb + object + *into* + gerund-participle,' as in *I talked Joe into unlocking the door*, and is provided by Anna Wierzbicka (1998). Since a number of her claims are problematic from the empirical point of view, it seems logical to start by examining them using modern electronic tools in order to clarify the nature of the data concerning these constructions before attempting to propose a hypothesis to explain them.

The first problematic claim in Wierzbicka's discussion is the following:

> In the case of the *into* construction, the causee originally didn't want to do what he or she did, whereas in the case of the *get* construction there is no such assumption. (Wierzbicka 1998: 125)

This claim is belied by the following attested use from COCA of the sequence 'verb + object + *into* + gerund-participle' in which the causees have presumably been longing to perform the event denoted by the gerund-participle for a considerable time:

(1) The premise is simple: Four unmarried women with love life "issues" are followed around by two love coaches (relationship columnist Teresa Strasser and JD Roberto, reality-show host of "Outback Jack"), who seek to guide them into finding the man of their dreams by giving them advice. (COCA)

Wierzbicka adds that "in the *into* construction, the causee's action is 'triggered' by the causer's will, not by the causee's own will, whereas in the *get* construction, the causee is acting in accordance with both his or her own will and the causer's will" (1998: 125–126). This claim is also problematic with respect to (1) above, in which the causees' will is part of the reason why they are aiming to perform the event denoted by the gerund-participle (*finding the man of their dreams*).

The inaccuracy of Wierzbicka's assertion is further confirmed by the following example from the Internet:

(2) He wanted to pursue painting after selling the company in 2000, and says a teacher "eased me into doing it full-time."
<www.culture24.org.uk/art/painting-and-drawing/art433683>

Here the person referred to is explicitly described as having wanted to do art full-time and was gradually able to achieve this goal under the teacher's guidance and encouragement.

A minor empirical problem concerns Wierzbicka's characterization of the '*get* + object + *to*-infinitive' construction as implying no assumption that the causee originally didn't want to do what he or she did. It should be pointed out to make things clearer that the *get* construction does not necessarily imply that the causee's will is involved at all, as can be seen from (3) below, which refers to a young baby:

(3) Basically, to get him back to bed I gave him 10 oz of formula and got him to burp a few times (it was pretty though, with him screaming).
(answers.yahoo.com)

Here the child evidently did not want to burp, and, in any case, burping is not a voluntary action for a young infant.

A second contention made by Wierzbicka that is not borne out by usage is that "in the *into* construction, the causee is unaware of what is happening (namely, that his or her action is 'triggered' by the causer's will), whereas in the *get* construction, there is no such assumption" (1998: 126). The incorrectness of this claim is demonstrated by examples (1) and (2) above in which the causee expressly wishes to perform the action denoted by the gerund-participle and is aware that his or her performance of this action is due to the causer.

Wierzbicka makes another empirical claim that is problematic when she states that "one cannot 'encourage' or 'induce' someone into doing anything" (Wierzbicka 1998: 125). Examination of real usage data in the BOE shows however that the sequences '*encourage* + *into* + gerund-participle' and '*induce* + *into* + gerund-participle' are in fact both attested:

(4) The new PRSA legislation is aimed at encouraging more people into making private provision for their retirement. (BOE)

(5) I am fairly confident in my bowling now – I try to induce batsmen into making mistakes. (BOE)

Another author who has dealt with the two patterns that are the object of our attention here is Juhani Rudanko, whose use of a corpus-based approach gives a sounder empirical basis to his observations. Thus he offers the very pertinent

consideration that "the *into -ing* pattern places an emphasis on the extent of the movement from the initial state to the final state and also carries the idea that the movement may have been by stages" (Rudanko 2006: 328). This observation can be illustrated by (6) below, which Rudanko cites in support of his analysis:

(6) You were mirroring a phenomenon that was happening in many gay lives across America during that time – an explosion of self-invention. Did this somehow lead you into studying the Native American *berdache*?
(BOE, cited by Rudanko 2006: 322)

Support for Rudanko's intuition can also be found in the fact that if the *to* + infinitive construction were used in the context of (6) above, the impression of a gradual waxing of interest in *berdache* would disappear and only the resultant impact on the interviewee's area of study would be evoked. Further support of his description is provided by the fact that the '*to* + infinitive' construction would not be compatible with the notion of gradual painless transition denoted by the verb *ease* in a context such as (2) above.

A second observation by Rudanko also hits the mark, this time regarding the infinitival construction: "The *to*-infinitive pattern focuses more on the resulting state, regardless of how it may be reached" (Rudanko 2006: 328). This remark is supported by the absence of the *into* + *-ing* construction with the verb *cause* (cf. Rudanko 2006: 319): since this verb presents the complement's event as a mere effect brought about by a cause, there is no concern with the means used to achieve the resultant effect. Further confirmation of Rudanko's observation is provided by the verb *get*, which tends to focus very strongly on the mere idea of achieving a result, as illustrated by uses such as *She got the job* and *We finally got to the top of the mountain*. While this verb is minimally compatible with the *into* + *-ing* construction, usage is preponderantly in favour of the *to*-infinitive structure. A search in the Corpus of Contemporary American English for the strings '*get* [verb] + *them*/*her* + *to* + infinitive' obtained 2468 attestations, while the corresponding sequence with '*into* + *-ing*' only found 3.

Another corpus-based study of the '*into* + *-ing*' construction, by Gries and Stefanowitsch (2004), provides a further relevant observation concerning this sequence:

> The verb most strongly associated with the [*into*] construction is *trick*, followed by *fool*, *coerce*, and *force*. (Gries & Stefanowitsch 2004: 227)

This statement is based on a measure of "collostructional strength," a calculation of association strengths between items performed on the basis of "a cross-tabulation of the individual frequencies of the word and the construction in question as well as their joint frequency" (Gries & Stefanowitsch 2004: 227). This measure places

trick ahead of *talk*, which is the most frequent collocate of the *into* construction in terms of raw frequency (in COCA, *talk* occurs 887 times followed by *into* + *-ing*, as compared to 536 for *trick*). One point of interest with respect to the above observation concerning the verb *force* is that, although this verb is high on the collostructional-strength scale with respect to other verbs in its association with the *into* + *-ing* construction, it is much more strongly associated with the *to* + infinitive structure than with *into* + *-ing*: a search in COCA for the strings '*force* + *them*/*her* + *to* + infinitive' obtained 2113 attestations, while the corresponding sequence with '*into* + *-ing*' only found 19. This distinguishes it sharply from its semantic neighbour *coerce*, of which 21 occurrences of the string '*coerce* [verb] + *them*/*her* + *into* + *-ing*' were found in COCA versus only 10 of the sequence '*coerce* [verb] + *them*/*her* + *to* + infinitive'.

3. Explanations anyone?

Although the three authors discussed above make many valid empirical observations and generalizations about the two structures under study, no explanations are proposed as to the origin of the semantic intuitions and statistical observations concerning these two causative constructions. That is what I will attempt to do in the rest of this article. The proposal put forward here will argue that the sequences in question do not constitute "constructions" in the sense of Goldberg (1995) in that observed usage is fully predictable from the semantic content of the items combined in these sequences.

The first question that we will attempt to answer is why the '*into* + *-ing*' construction tends to suggest "manipulation" (Wierzbicka 1998) or "trickery" (Gries & Stefanowitsch 2004). In our view, this is due mainly to the meaning of the preposition *into*, which denotes movement leading to containment. This can be illustrated by its spatial use in a context such as (7):

(7) He walked into the room.

The notion of movement leading to containment signified by this preposition holds the potential for conveying the impression of entrapment, i.e. of maneuvering someone into a situation that they did not want to be in. In this respect, it is significant that the second most frequent pair of verbs found with this construction after *trick* and *fool* is *force* and *coerce*, both of which imply that the person denoted by the direct object does not perform the action expressed by the gerund-participle willingly. This impression is also observed with the verb *induce* in (5) above. On the grammatical level, the gerund-participle's grammatical meaning also contributes to the expression of this message: Duffley (2006: 19–20) argues

that the schematic meaning of the gerund-participle puts the implicit subject of the *-ing* form in relation with the interiority of the event. This is most evident in the progressive, where the subject of the auxiliary *be* is situated within the interiority of the event expressed by the gerund-participle at the moment in time evoked by the auxiliary, as in:

(8) When I saw her, she was playing tennis with her sister.

In its gerundive use, the gerund-participle evokes the event's interiority as a homogenous whole made up of all of the instants contained between its beginning and its end points, a construal which is aspectually neutral (see Duffley 2014: 35–36, De Smet 2010: 1169) and can denote both an event that is completely accomplished (as in *I remember playing tennis with her sister*) and one that is in progress, as in the use below involving the preposition *in*:

(9) She was so engrossed in playing tennis with her sister that she didn't even notice my arrival.

In the *into* + *-ing* construction, the object of the main verb ends up being involved in the event denoted by the gerund-participle, as illustrated by the pair of sentences in (10)–(11), the second of which could describe the situation immediately ensuing upon that depicted by the first:

(10) Nita is a huge fan of Jason Becker, and her boyfriend/manager tricked her into playing a Jason Becker song for her hero!
(wgrd.com/alice-cooper-guitarist-nita-strauss-was-tricked-into-playing)

(11) She was engaged in playing the Jason Becker song when I walked in.

Thus both the meaning of the preposition *into* and the way the gerund-participle construes the event that it denotes make a contribution to the effect of entrapment and the attendant impressions of manipulation or trickery observed by Wierzbicka and Gries & Stefanowitsch. It should be noted, however, that this impression is not present in all occurrences of this construction:

(12) She finally wore me down and talked me into buying a package of cigarettes.
<https://books.google.ca/books?isbn=149076173X>

In this use, the main idea is that of pushing the person into an action that they did not want to perform.

The next question to be addressed is the explanation of Rudanko (2006)'s observation that the '*into* + *-ing*' pattern places an emphasis on the extent of the movement from the initial to the final state and gives the impression that the movement may have been by stages. It will be argued here that this is also principally due to the meaning of the preposition *into*, which represents the movement evoked by

to as leading to the complete containment of the trajector within the landmark. In his study of the distinction between *in* and *into*, Tutton (2009: 20) notes that "*into* seems to accentuate path to a greater extent than boundary-crossing *in*." This is due to the *to*-component of its meaning. Due to its *in*-component, *into* adds to the extension of the movement leading up to the container the additional stages of penetrating the container's external limit and getting all of the trajector inside. This is illustrated in Figure 1:

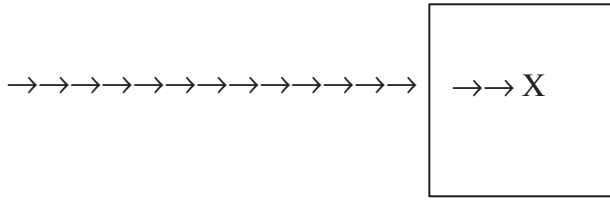

Figure 1. *Into* as movement leading to containment

It is revealing to compare the impression produced by the '*into* + *-ing*' construction with that observed with '*to* + infinitive'. Rudanko remarks quite judiciously that the *to*-infinitive pattern focuses more on the resulting state, regardless of how it may be reached, but does not propose any explanation as to why this should be so. It will be held here that this is due to the meaning of the components of this construction, the preposition *to* and the bare infinitive. The preposition represents the terminus of the movement that it denotes as a point rather than a container, as illustrated in Figure 2:

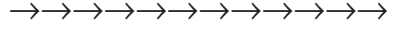

Figure 2. *To* as movement leading to a point

The bare infinitive, for its part, denotes the complete actualization of its event in time (cf. Duffley 2006: 28–30), or in Langacker's terms (1987: 250–252), the "full instantiation" of the event. This semantic configuration focuses on the mere achievement of the infinitive event's full actualization as a result of the matrix verb's event. There is no notion of penetration of an external limit or entry into the interior of a container, and so no impression of stages in the movement leading to the complement event's realization.

This account also provides an explanation of why the verb *ease* is not found with the '*to* + infinitive' construction. Since this verb denotes a gradual, painless transition into a new state, it is not semantically compatible with the notion of the mere achievement of resultant realization expressed by the '*to*

+ infinitive' pattern. Encouragement, on the other hand, can be construed as gradually coaxing someone into a new way of doing things, and this explains why the '*encourage* + *into* + gerund-participle' pattern is attested in uses like (4) above. In addition, an explanation can also be offered for the fact that a verb like *talk* is not compatible with the '*to* + infinitive' construction. The reason is quite simple: talking is not goal- or result-oriented. On the other hand, the verb *convince*, which is very clearly result-oriented and focuses on obtaining the adhesion of the convincee to the opinion of the convincer, is almost exclusively construed with *to* + infinitive: only 5 *into*-constructions are attested in the 553-million-word Bank of English corpus vs 6851 attestations with the infinitive, and only one occurrence of this pattern is found in the 450-million-word Corpus of Contemporary American English, and that in a parallel construction with the verb *force*:

(13) A few of the young women reported that they had been more forcibly coerced into sexual intercourse, that is, that their boyfriends had used pressure to *convince or force* them into having sexual contact against their will. (COCA)

The *to*-infinitive pattern, on the other hand, is attested 3718 times with the verb *convince* in COCA. *Persuade*, in comparison, is more amenable to the '*into* + *-ing*' construction: there are 22 occurrences of this pattern in the Bank of English (cf. Hunston & Francis 2000: 103) and 12 in COCA (versus 4288 of the *to*-infinitive structure). This is because *convince* is result-oriented while persuasion involves both process and result. *Convince* derives from the Latin verb meaning 'conquer, overcome'; to *convince* someone of something is to cause that person to believe the truth of something by means of facts, arguments, etc.; *persuade* comes from the Latin 'advise, urge until the attainment of a result' and can be defined as 'to bring (someone) to do or believe something through reasoning, argument or appeal to emotions'. Further evidence of the result-focussedness of *convince* versus the inclusion of the idea of process with *persuade* is found in two facts of usage. The first is that if we are absolutely sure about something we say *I'm convinced* rather than *I'm persuaded*:

(14) – Are you sure he's guilty?
– Yes, I'm convinced.

Here the speaker wants to express solely the resultative state of mind that he is in with respect to the person's guiltiness. The second fact is that one can say *to use persuasion on someone*, but not **to use conviction on someone*: COCA has 22 occurrences of the former sequence and none at all of the latter. Here persuasion

is construed as the means employed in order to get someone to do something. The subtle semantic difference between these two verbs is thus reflected in their relative frequency with the '*into* + gerund-participle' pattern.

An explanation can also be proposed for the constructional difference noted above between *force* and *coerce* based on the lexical meanings of these two verbs. The fact that *coerce* is more frequent with the *into -ing* structure than with the *to*-infinitive is a reflection of the original ideal of hemming someone in (*co-* + *arcere* 'to shut up, enclose') expressed by this verb, which according to the OED involves "the application of force to control the action of a voluntary agent" and according to Webster's places emphasis on "domineering and overriding resistance": to coerce is to forcibly constrain someone so that they are funneled into a situation against their will. The verb *force*, for its part, evokes the exercise of the power required to move an object and so is goal- or result-oriented, which explains its marked preference for the *to*-infinitive. This aspect of its meaning is also reflected in its occurrence in structures such as *to force a door open* in which the adjectival object complement denotes the resultant state into which the door is put by means of the exercise of force.

Table 1 provides a summary of the comparative frequencies of the sequences examined thus far:

Table 1. Summary of comparative frequencies

Sequence	Frequency
get + them/her + to-infinitive (COCA)	2468
get + them/her + into + -ing (COCA)	3
talk + NP + into + -ing (COCA)	887
trick + NP + into + -ing (COCA)	536
force + them/her + to-infinitive (COCA)	2113
force + them/her + into + -ing (COCA)	19
coerce + them/her + to-infinitive (COCA)	10
coerce + them/her + into + -ing (COCA)	21
convince + NP + to-infinitive (COCA)	3718
convince + NP + into + -ing (COCA)	1
convince + NP + to-infinitive (BOE)	6851
convince + NP + into + -ing (BOE)	5
persuade + NP + into + -ing (BOE)	22
persuade + NP + into + -ing (COCA)	12
persuade + NP + to-infinitive (COCA)	4288

4. Construction grammar, embodied cognition and the basic design architecture of human language

George Lakoff and his co-researchers (Lakoff 1987; Lakoff & Johnson 1980, 1999; Lakoff & Turner 1989) have developed several lines of evidence in support of the hypothesis that language involves embodied cognition, i.e. that people use their understanding of familiar physical objects, actions and situations such as containers, spaces and trajectories to understand other more complex domains. The data examined in this study provides further evidence in favour of this hypothesis, as it shows that causing someone to do something by means of convincing is construed in terms of a trajectory leading to a point, while causing someone to do something by means of talking is conceived in terms of a trajectory leading to containment. There is, however, another dimension of embodiment relevant to our discussion that is often overlooked, and that has to do with the basic design architecture of human language itself: Ronald Langacker (1987: 11) notes the fundamental fact that "language is symbolic in nature," by which he means that the foundational relation on which human language is based is the association between a mind-engendered meaning and a bodily-produced sign. It is our contention that many of the problems that we have pointed out in Wierzbicka's, Rudanko's and Gries & Stefanowitsch's analyses stem from a neglect of this basic principle.

A large part of the cause of this oversight stems from treating meaning on the level of the overall construction rather than of the linguistic signs of which the construction is composed. This approach is characteristic of Construction Grammar, which works on the assumption that: (1) "constructions are taken to be the basic units of language" (Goldberg 1995: 4) and (2) "constructions are independent of the lexical items which instantiate them" (Goldberg 1995: 1, 4). Regarding the first postulate, it should be remarked that it only makes sense if one follows Goldberg in expanding the notion of 'construction' to cover morphemes, based on the fact that the latter are "pairings of meaning and form that are not predictable from anything else" (Goldberg 1995: 4). This, however, amounts to an abuse of the term 'construction,' as on the syntactic level this term corresponds to a combination of items that all have meanings, whereas on the morphemic level it denotes the combination of a meaningful meaning with an as yet meaningless form. As for the second assumption, it needs to be handled with care as it has the potential for introducing a certain disconnection between linguistic meaning and linguistic form. This occurs, for example, in Wierzbicka's analysis, where the meanings attributed to the '*get* + object + *to* + infinitive' and 'verb + object + *into* + gerund-participle' constructions are much too specific and consequently unable to cover all of the uses of the two patterns in question. There are two main reasons for this. The first is that in the majority

of cases constructions are discourse entities, assembled on-line for specific communicative purposes.[1] Since the most salient aspect of discourse entities is the specific message that they serve to communicate and not the linguistic means used to convey this message, there is a high risk of confusing linguistic meaning with the specific message intended when one analyses meaning on the level of the construction. The intended message, however, is the end-product of an inference that the speaker wants the hearer to make, and is based not just on the linguistic meanings contained in the utterance but also on the utterance situation, on what has been said previously by the interlocutors, on shared knowledge of the world, etc.

In Wierzbicka's case, the use of Natural Semantic Metalanguage as a tool for describing meaning constitutes a second confounding factor in the analysis, as it distorts the description of meaning by forcing the latter into the mold of a limited set of universal semantic primitives, a procedure which is somewhat analogous to translating the English constructions into a foreign language in order to describe their import. This compounds the confusion between linguistic meaning and intended message, as the equivalence between the Natural Semantic Metalanguage paraphrase and the original English construction lies on the level of the message and not on that of the linguistic means used to convey that message (cf. the equivalence in the message communicated between the Spanish utterance *Tengo hambre* and the English *I am hungry*, where there is no notion of 'tener' or 'having' in the English utterance, and no notion of 'being' in the Spanish, in addition to the conceptualization of *hambre* as a noun in a direct-object relation to the verb *tener* in the Spanish sentence vs. the construal of *hungry* as an adjective in subject complement function in the English one).

Dissociating linguistic meaning from linguistic form represents a failure to respect the principle of embodiment as it applies to human language on the most basic level, namely the fact that language is symbolic in nature. Generative Grammar disregards this principle by attempting to deal with form in abstraction from meaning on the syntactic level. Chomsky (1957: 141) famously argued that since meaning is "notoriously difficult to pin down," if it were to be shown that it played a central role in linguistic analysis the latter's results and conclusions would "become subject to all the doubts and obscurities that plague the study of meaning," thereby striking a serious blow at the very foundations of linguistic theory. In his view it is preferable therefore to set meaning aside and focus on formalizing the distributional configurations of the physically observable linguistic signs that make up the

1. This has been known for a long time: Hermann Paul observed in 1880 that "one simply has to admit that only very few sentences have been memorized as such. Most sentences are composed on the spur of the moment" (pp. 109–110).

sentence, i.e. to build a formal syntax of the language. Since, however, meaning cannot be kept out of the picture completely, as it is obvious that people speak in order to express their ideas, it must be re-introduced into the analysis at some point. In a generative model, this is done only after the syntactic component has generated a string of linguistic forms, which is then sent to the semantic component for interpretation. This theoretical model raises a number of important problems. Firstly, as pointed out by Seuren (2004: 161), there is the crucial question of what drives the syntactic sentence generator. To claim that the generation of a sentence is "a process, activated by a start signal, that randomly selects lexical items and casts them into a grammatical structure" is absurd according to Seuren. Equally absurd is "the notion that a randomly generated sentence structure should be taken to pass an instruction to the cognitive system of the same organism for the sentence to be interpreted": this would entail that the speaker does not have any idea in mind to express before he starts generating a syntactic string of linguistic forms. Serious problems thus arise when form is treated autonomously from meaning.

The opposite pitfall also exists, however, i.e. treating meaning in abstraction from form. This is a significant risk for approaches that situate meaning on the level of the utterance or the construction such as that of Evans (2009: 25), who holds that "words do not in fact have meaning (…). In my account, meaning is a function of an utterance, rather than a given lexical representation associated with a word or other symbolic (i.e. linguistic) unit." The problem with this type of account is that the relation between message and utterance/construction is not usually stable: some sentence- or construction-level assemblages clearly do have a stable semantic content (e.g. *All that glitters is not gold*; *to cut the mustard*); in most cases, however, that is not the case. Thus, as noted by Bardzokas (2012: 29), in an utterance of the sequence *John is too tired* the speaker is not understood to be simply predicating of John the property of excessive tiredness, but to be putting his degree of tiredness into relation with some action that John might be expected to perform in the situation to which the utterance refers. The message conveyed thus varies in function of what it is that John is too tired to do. Correlating meaning and form on the level of the utterance leads therefore to a proliferation of different meanings for the same form. This is exemplified in Evans' analysis of the adjective *long* as having "at least two conventionally established lexical concepts" (p. 299) – [EXTENDED IN HORIZONTAL SPACE], as in *a long stick*, and [EXTENDED IN DURATION], as in *a long kiss*. This amounts, however, to treating as part of the adjective's lexical concept something which is contributed by the lexical concept of the noun following it.

Something similar occurs in some applications of Construction Grammar. A case in point is Stefanowitch (2001)'s construction-grammar approach to English analytic causatives. This author posits three "causation event types" as being relevant to these constructions: the MANIPULATE type, where an animate causer

intentionally acts on a causee in a way that influences the causee such that he or she performs some activity (e.g. *This guy was taking women from the teller and making them give him money*); the TRIGGER type, where an event occurs that influences the causee in such a way that the latter inevitably undergoes some process (e.g. *Seeing people abuse the system makes me feel bad*); and the PROMPT type, where an event occurs and the causee perceives this event and decides to react by performing some activity (e.g. *What made her decide to go into an old people's home?*). Stefanowitch himself observes (2001: 311) that the MANIPULATE, TRIGGER and PROMPT configurations "are not tied to any particular constructions," but occur across structures with the four main causative verbs *make*, *have*, *get* and *force*. In the illustrative examples just given above, the very same verb *make* occurs in all three types. This entails that *make* is in fact indifferent to Stefanowitch's causation event types and must be treated as having a more abstract type of meaning that has the potential to correspond to all three of them. With regard to the two constructions examined in the present study the same thing is true. Both *talk into* and *convince to* can correspond to the MANIPULATE type:

(15) Hello can someone give me some tips? For a third day straight this guy (friend's friend) talked me into giving him some money.
<https://answers.yahoo.com>

(16) Some bum named Slim followed me around and convinced me to give him some money for a local bus to another town.
<https://sellersabroad.wordpress.com>

In addition, *convince* + *to*-infinitive can convey the notion of PROMPT:

(17) It bugged him for a while, but her reaction convinced him to drop it until he could sleuth it out on his own.
<http://apothica.forumotion.com/t6-apothic-black-figures-in-the-dark>

Such causation event types appear therefore to be tangential to the type of meaning expressed by the '*convince* + object + *to*-infinitive' and '*talk* + object + *into* + gerund-participle' constructions. This is not surprising, in that the event types in question were posited independently of any reference to linguistic sign-meaning units in the first place.

Another illustration of the metalinguistic character of some constructions can be found in Rudanko (2015)'s treatment of the transitive *into -ing* structure. Following on Goldberg (1995)'s general theory, he classifies this sequence as a "caused motion construction." This leads him, however, to exclude the following attestation from the scope of his study:

(18) Just as he once battled for supreme fitness, he has poured his energy into learning to speak again. (BNC)

The reasons given for this exclusion are that "the direct object designates a resource at the disposal of the referent of the higher subject, and the pattern may be viewed as one of subject control." Regarding the allusion to subject versus object control, since it has been demonstrated that such distinctions are substantially pragmatic in nature (Duffley 2014), one wonders whether they can be validly invoked as evidence of the existence of different grammatical constructions. More fundamentally, the question is raised by this Construction Grammar approach as to why *He charmed the child into learning to speak again* represents a construction, whereas (18) above does not. It is argued here that both are simply possible exploitations of the semantics and syntax of the Verb + NP + *into* + *-ing* sequence, and that the difference between them is merely a product of the differences in meaning between the verbs *charm* versus *pour*, the direct objects *the child* versus *his energy*, and the pragmatics of their interrelations.

A similar problem of the Construction Grammar approach leading to a reification of the meaning of constructions due to the fact that it postulates a stable form-meaning correlation on the level of the overall syntactic configuration is found in Kim and Davies (2015: 81), where the transitive *into -ing* sequence is once again treated as a "caused motion construction" distinguished from other such constructions by the fact that "it entails that the situation denoted by the gerundive phrase actually happened." While this is usually the case with the sequence in question, nevertheless it depends crucially on the semantic content of the matrix verb. Thus just as *Mary invited the woman into her house* does not necessarily entail that the woman actually came into her house, the sentence below does not necessarily entail that the addressee actually explores what may keep them from love:

(19) When you stop to listen, you hear a small but persistent voice that urges you towards healing. This voice invites you into exploring what may keep you from love, belonging, and satisfying relationships.
<www.tamaralynnrobert.com>

According to Kim and Davies' criterion, this means that some instances of the sequence *invite* + NP + *into* + *-ing* should be treated as caused motion constructions and others not. Such a division seems artificial and non-linguistic.

The general methodological point that I wish to make here is that one should first investigate whether the message conveyed by a construction can be explained by the stable form/meaning units of which it is composed before attributing a meaning to the whole construction or attempting to fit it into a set of sign-independent notional categories. The case of Wierzbicka's analysis of the '*get* + object + *to* + infinitive' and 'verb + object + *into* + gerund- participle' constructions examined here is an example of 'jumping the gun,' where the analysis skips over the level where the form-meaning relation is stable to the level of the

utterance-message relation, where it is not. Stefanowitch's approach, on the other hand, starts with abstract non-symbolic categories and works back towards the linguistic forms, but does not succeed in connecting the latter to these categories in any direct or stable manner. This paper has attempted to show the explanatory gain that can be achieved by starting one's analysis on the level on which meaning is stably embodied, which is normally that of the word or morpheme, where a linguistic sign is stored in a stable, permanent and direct relation with its meaning outside of any particular context.

References

Aissen, Judith. 1979. *The Syntax of Causative Constructions*. New York NY: Garland.

Baron, Naomi S. 1974. The structure of English causatives. *Lingua* 33: 299–342. doi:10.1016/0024-3841(74)90044-8

Bardzokas, Valandis. 2012. *Causality and Connectives. From Grice to Relevance* [Pragmatics & Beyond New Series 216]. Amsterdam: John Benjamins. doi:10.1075/pbns.216

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Comrie, Bernard & Polinsky, Maria. 1993. *Causatives and Transitivity* [Studies in Language Companion Series 23]. Amsterdam: John Benjamins. doi:10.1075/slcs.23

De Smet, Hendrik. 2010. English *-ing* clauses and their problems: The structure of grammatical categories. *Linguistics* 48: 1153–1193. doi:10.1515/ling.2010.038

Duffley, Patrick J. 2006. *The English Gerund-participle*. A Comparison with the Infinitive. Frankfurt: Peter Lang.

Duffley, Patrick J. 2014. *Reclaiming Control as a Semantic and Pragmatic Phenomenon* [Pragmatics & Beyond New Series 251]. Amsterdam: John Benjamins.

Evans, Vyvyan. 2009. *How Words Mean. Lexical Concepts, Cognitive Models, and Meaning Construction*. Oxford: OUP. doi:10.1093/acprof:oso/9780199234660.001.0001

Gilquin, Gaëtanelle. 2010. *Corpus, Cognition and Causative Constructions* [Studies in Corpus Linguistics 39]. Amsterdam: John Benjamins. doi:10.1075/scl.39

Givón, Talmy. 1975. Cause and control: On the semantics of interpersonal manipulation. In *Syntax and Semantics*, Vol. 4, John P. Kimball (ed.), 59–89. New York NY: Academic Press.

Givón, Talmy. 1980. The Binding Hierarchy and the typology of complements. *Studies in Language* 4: 333–377. doi:10.1075/sl.4.3.03giv

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Gries, Stefan T. & Stefanowitsch, Anatol. 2004. Covarying collexemes in the *Into*-causative. In *Language, Culture, and Mind*, Michel Achard & Suzanne Kemmer (eds), 225–236. Stanford CA: CSLI.

Hollmann, Willem B. 2005. The iconicity of complementation in present-day English causatives. In *Outside-in–inside-out* [Iconicity in Language and Literature 4], Constantino Maeder, Olga Fischer & William J. Herlofsky (eds), 287–306. Amsterdam: John Benjamins. doi:10.1075/ill.4.21hol

Hunston, Susan & Francis, Gill. 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English* [Studies in Corpus Linguistics 4]. Amsterdam: John Benjamins. doi:10.1075/scl.4

Kim, Jong Bok & Davies, Mark A. 2015. The *into*-causative construction in English: A construction-based perspective. *English Language and Linguistics* 20: 55–83.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago IL: University of Chicago Press. doi: 10.7208/chicago/9780226471013.001.0001

Lakoff, George & Johnson, Mark. 1980. *Metaphors We Live By*. Chicago IL: University of Chicago Press.

Lakoff, George & Johnson, Mark. 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York NY: Basic Books.

Lakoff, George & Turner, Mark. 1989. *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago IL: University of Chicago Press. doi: 10.7208/chicago/9780226470986.001.0001

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. 1: Theoretical Prerequisites. Stanford CA: Stanford University Press.

Paul, Hermann. 1880[1920]. *Prinzipien der Sprachgeschichte*. Halle: Max Niemeyer.

Rudanko, Juhani. 2006. Emergent alternation in complement selection. The spread of transitive *into -ing* constructions in British and American English. *Journal of English Linguistics* 34: 312–331. doi: 10.1177/0075424206293600

Rudanko, Juhani. 2015. *Wheedled me into lending him my best hunter*: comparing the emergence of the transitive *into -ing* pattern in British and American English. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 128–140. Houndmills: Palgrave Macmillan.

Seuren, Pieter A. M. 2004. *Chomsky's Minimalism*. Oxford: OUP.

Shibatani, Masayoshi. 1976. The grammar of causative constructions. In *Syntax and Semantics*, Vol. 6: *The Grammar of Causative Constructions*, Masayoshi Shibatani (ed.), 261–312. New York: Academic Press.

Song, Jae Jung. 1996. *Causatives and Causation. A Universal-typological Perspective*. London: Addison Wesley Longman.

Talmy, Leonard. 1976. Semantic causative types. In *Syntax and Semantics*, Vol. 6: *The Grammar of Causative Constructions*, Masayoshi Shibatani (ed.), 43–116. New York NY: Academic Press.

Tutton, Mark. 2009. When *in* means *into*: Towards an understanding of boundary-crossing *in*. *Journal of English Linguistics* 37: 5–27. doi: 10.1177/0075424208329308

Verhagen, Arie & Kemmer, Suzanne. 1997. Interaction and causation: causative constructions in Modern Standard Dutch. *Journal of Pragmatics* 27: 61–82. doi: 10.1016/S0378-2166(96)00003-3

Wierzbicka, Anna. 1998. The semantics of English causative constructions in a universal-typological perspective. In *The New Psychology of Language*, Vol. 1, Michael Tomasello (ed.), 113–153. Mahwah NJ: Lawrence Erlbaum Associates.

Passive permissives

Being let and allowed

Thomas Egan
Hedmark University of Applied Sciences

The '*let* NP bare infinitive' construction differs from other common permissive constructions, such as '*allow* NP *to*-infinitive' and 'permit NP *to*-infinitive' in being exceedingly rare in the passive. That is, while somebody may well be 'allowed to do' something, they are very seldom 'let do' something. Even more seldom are they 'let to do' something. This chapter explores possible reasons for the rarity of both of these passive *let* constructions, which are contrasted with passive *allow* constructions. It is argued that the difference in distribution between the constructions with the two matrix verbs is related to two factors. The first is a difference in the sorts of force dynamic relations which they typically encode. The second is related to the difference in semantics between the two infinitive forms. The argument is supported by corpus data from both British and American English.

Keywords: permissives; causatives; force dynamics; enablement

1. Introduction

The matrix verbs *let* and *allow* are both very common in English. Indeed, when used to report permission in active clauses they are about equally common (Egan 2008: 220). When the matrix verb is in the passive voice the situation is very different, however. While *allow* is again one of the half dozen most common matrix verbs in English, *let* is extremely rare, being represented by only 22 tokens in the whole of the British National Corpus (BNC), as shown in Table 1.

Table 1. Tokens of active and passive *allow* and active *let* in the BNC, projected on the basis of a randomly downloaded sample of 1,000 tokens, with the actual total number of tokens of passive *let*

	***allow to* V**	***let* V**
active matrix verb	15,300	14,100
passive matrix verb	4,700	22

DOI 10.1075/slcs.195.03ega

Thus the construction in (1) is very common, the construction in (2) very uncommon.

(1) Few towns can boast they do not have a problem with dogs *being allowed to roam* the streets and Darlington is no exception. (BNC K54 6237)

(2) She *shouldn't be let roam* the hills alone. (BNC G0X 7)

The fact that *let* tends to be avoided in the passive has often been noted. Thomson and Martinet (1986: 23), for instance, note that "*let* in the passive is often replaced by another verb", and Carter and McCarthy (2006: 99) agree that "*let* is not normally used in the passive when it means 'allow/permit'". However, there is very little in the literature about why *let* should resist the passive. After all, other verbs that occur with bare infinitive complements, such as *make* and perception verbs, do occur in the passive, albeit with a *to*-infinitive form of complement. Indeed '*be seen to* V' and '*be made to* V' are among the 25 most common constructions containing passive matrix verbs and non-finite complements (Egan 2003: 415). In this chapter I examine both active and passive complement clauses containing *let* and *allow* from the point of view of force dynamics. I show that the almost complete absence of *let* passives is at least in part a consequence of the type of force dynamic relations prototypically encoded by *let*.

This chapter is structured as follows. Section 2 contains an analysis of the force dynamic relations of *let* and *allow* constructions with both positive and negative active matrix verbs in the BNC. Section 3 contains similar data and analysis of constructions with passive matrix verbs. Section 4 discusses data from the Corpus of Contemporary American English (COCA), including examples of the passive *let to*-infinitive construction, which is not found in the BNC (apart from a quotation from *The Origin of Species*). Section 5 considers the development of both passive *let* constructions in the light of the evidence of the Corpus of Historical American English (COHA). Finally, Section 6 contains a summary and conclusion.

2. Constructions with active matrix verbs[1]

The data for Sections 2 and 3 were taken from the BNC. As there are almost 30,000 tokens of the verb *let* and over 30,000 of *allow* in the BNC, it was impossible within the scope of this study to investigate all tokens. I decided to restrict my investigation to a certain number of randomly chosen utterances containing each matrix

1. For a more detailed analysis of the constructions in this section see Egan (2008: 214–237).

verb, and to examine this subset of utterances for tokens containing bare and *to*-infinitive complement clauses. 1,000 randomly selected tokens of *let* and *allow* were downloaded from the corpus and the tokens containing non-finite complements were extracted from the two databases. There were in all 621 tokens of *allow* and 774 of *let* with non-finite complements in the downloaded sample. Both *allow* and *let* are polysemous and both may be used in senses that do not encode permission (or prohibition) as such. It was therefore necessary to weed out from the data tokens which do not encode permission. These include (3) in which *allow* means *admit* or *consider* rather than *permit*. Also excluded were tokens of the two multiword verbs '*let* x *know*' (= 'inform x') and '*let* x *have*' (= 'give x'), illustrated in (4) and (5), and first-person plural suggestions, of the sort illustrated in (6).

(3) We can apply the test to the technical and technological subjects, and not only those, but the professional subjects also; and the boundary line will run now on this side, now on that; but the things that it divides are different in kind, and only on one side of that line lies what we *ought to allow to be* education. (BNC A69 383)

(4) And we *let her know* from the start that we trusted her. (BNC G35 1029)

(5) I *will let you have* a list of his customers and I want them contacted, in the first instance by telephone. (BNC HWP 1159)

(6) *Let's assume* one of your employees drinks too much both at work and at home. (BNC A05 29)

Examples (4) and (5) instantiate causative rather than permissive constructions, coding mental and physical transfer respectively. Another causative construction which was omitted from the study is 'let x go' meaning 'fire x', as in (7). On the other hand, I decided to include the (at least partially) lexicalised permissive '*let* x *go*' meaning 'release x', as in (8) and (9). There are three reasons for retaining this construction. In the first place it is clearly permissive, rather than causative, at least in cases where x is animate. The second reason is that if '*let* x *go*' were to be considered completely lexicalised, the '*let* x *go free*' construction, instantiated in (9), would be tautological. The third reason is that '*let* x *go*' may encode various types of permissive event which do not involve releasing, as in (10), (11) and (12).

(7) We had a drummer and he just couldn't cut it. We *had to let him go*. (BNC CKS 776)

(8) I beg those who are holding her *to let her go*. (BNC G0N 1517)

(9) The police had found somebody who had seen me on my island at the time of the murder, and so the judge *let me go free*. (BNC H8G 649)

(10) When we get back to the house, Marie gives me the key to the door and I unlock it and *let her go* in first. (BNC A74 1184)

(11) The staff thought it better to *let her go* through her labour cheerfully. (BNC G0T 643)

(12) If it comes out of your mouth spit it out, rather than *let it go* back down to your tummy. (BNC KBW 5219)

Having removed the non-permissive tokens we are left with 462 tokens of active voice '*allow* x *to*-infinitive' and 490 tokens of '*let* infinitive' that clearly encode permission or its negative counterpart, prohibition. There were also 141 relevant tokens of '*be allowed to*-infinitive'. These will be discussed in Section 3. There was not a single token of '*be let* bare infinitive' among the 1,000 randomly downloaded tokens of *let*.

In an influential paper on causation, Kemmer and Verhagen (1994) characterise permissives as encoding the removal of a barrier preventing what I will term the 'permittee' from realising some goal.

> A fourth type [of causation], *enablement/permission*, involves not the exertion of force on an entity to bring about an event that otherwise would not have happened, but the removal by the causer of a conceived barrier that was preventing the causee from carrying out or undergoing the effected event. *Enablement* refers to the case where the barrier is physical […] and *permission* to the case where the barrier is social or sociopolitical in nature […]; we can thus consider enablement and permission as subvarieties of one type. (Kemmer & Verhagen 1994: 120)

Figure 1 illustrates this type of permission or enablement, wherein the matrix verb subject, the permitter (S1, which stands for first subject), removes a barrier which was blocking the path of the complement verb subject, the permittee (S2, or second subject), permitting the latter to continue unimpeded on his or her way.

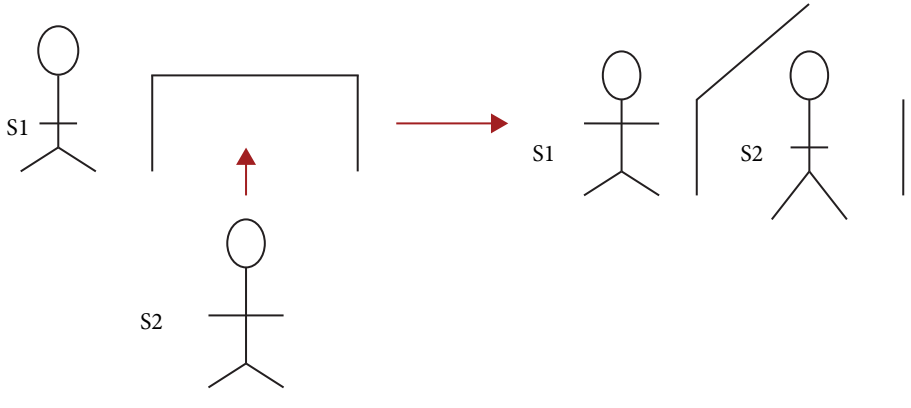

Figure 1. *Barrier-removal* by the permitter (S1) enabling the permittee (S2) to pass

Figure 1, however, illustrates only one of two main forms of permission described by Talmy (1986), who distinguishes between what he calls *onset letting* and *extended letting* as follows: 'onset letting correlates with the cessation of impingement and extended […] with its nonoccurrence' (Talmy 1986: 76, see also Talmy 2000: 418). While accepting Talmy's distinction between these two types of permission, I prefer to use the term *barrier-removal*, based on Kemmer and Verhagen (1994), rather than *onset-letting*. For the concept which Talmy calls *extended letting* I will use the term *non-imposition* (of any barrier). I will also eschew Talmy's terminology (*agonist* and *antagonist*) for the participants in the act of permission, preferring the more specific terms *permitter* and *permittee*. The form of permission which I term *non-imposition* is illustrated in Figure 2.

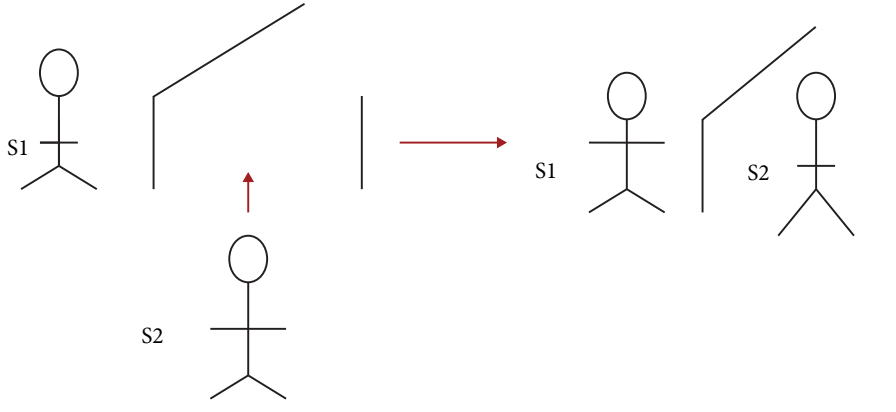

Figure 2. *Non-imposition* of barrier by the permitter (S1) enables the permittee (S2) to pass

Figures 1 and 2 illustrate permission rather than its opposite, prohibition. It is only matrix verbs with positive polarity that encode *barrier-removal* or *non-imposition*. Negative polarity matrix verbs encode *barrier-retention* or *imposition*. These will be discussed below. Table 2 contains details of the numbers of positive and negative polarity matrix verbs in the downloaded samples.

Table 2. Constructions containing positive and negative polarity active voice matrix verbs *allow* and *let* with horizontal percentages

	Totals		Percentage totals	
Matrix verb	**positive**	**negative**	**positive**	**negative**
allow	414	48	89.6%	10.4%
let	372	118	76.0%	24.0%

All positive polarity tokens were examined with a view to determining whether they encoded *barrier-removal* or *non-imposition*. The two types of permission were taken to comprise mutually exclusive categories – either a barrier existed or it did not. Distinguishing between the two sometimes involves a considerable amount of trawling through the co-text in an effort to ascertain the possible prior existence of barriers. In other cases the immediate co-text contains sufficient information to conclude that such a barrier existed. Possible evidence for the existence of a barrier may include the presence of a temporal adverbial like *later* in (13) or an adjective like *new* in (14).

(13) The US pilots later *allowed an Iraqi search-and-rescue helicopter to fly* to the crash site and then return to its base. (BNC CBE 784)

(14) In an attempt to remedy this the SLORC introduced new banking laws in July 1990 which *allowed foreign banks to open* branches in Myanma. (BNC HLD 4402)

We can also make inferences about the prior existence of a barrier on the basis of other sorts of information in the immediate co-text, as in (15), or using our general world knowledge as in (16).

(15) She *allowed herself to feel* all the pain she'd denied herself for so long. (BNC HGM 851)

(16) Claudia relaxed her fingers, *letting the pencil drop* to the desk. (BNC H8J 2708)

In (15) it is the presence of the adverbial *for so long* in the relative clause that allows us to infer the previous self-imposed barrier to the feeling of pain. In (16) our knowledge of the function of taut fingers as a container of objects allows us to conclude that prior to their being relaxed the fingers constituted a barrier to the pencil's falling.

Another type of barrier takes the form of a *sine qua non* condition, as in (17)–(18).

(17) If you re recall back in nineteen eight five Tony the Government brought in the transport bill which *let operators compete*. (BNC KM8 236)

(18) The two centre holes *allow a retaining wire to be fitted*. (BNC HH6 1902)

(17) is similar to example (13) in that it contains a temporal adverbial, *back in nineteen eight five*. However, the presence of the adverbial is not necessary for us to make the requisite inference. The very fact that it is the bill that is the permitter implies the prior impossibility of competition, in other words the existence of an earlier impediment. Similarly in (18) without the presence of the two centre holes

a wire could not have been fitted. Thus the presumed absence of these two holes amounts to a prior barrier.

Examples (13)–(18) all encode situations of *barrier-removal*. To categorise them as such it is sufficient to identify the earlier existence of a barrier, which may either be implicit or explicit. The prior non-existence of a barrier is less easy to stipulate, for obvious reasons. We may sometimes draw on our world knowledge, as in the case of (19)–(21). Often we must trawl the co-text before we can conclude that no such barrier existed.

(19) With the tension reaching boiling point, it was finally announced that the French officials *had allowed the result to stand* and they had to be applauded for a sporting decision. (BNC A40 42)

(20) Have they all *let their membership lapse*? (BNC HHV 24488)

(21) Race starter Captain Keith Brown was also criticised for *allowing the horses to line up* too close to the start line which led to the tape twice being broken. (BNC K45 1259)

(22) So we *let the blacks come down to us*, we didn't go looking for them. (BNC FAY 933)

We can infer from (19), without searching the co-text, that the officials in question had the power to alter the result but chose not to exercise this power. In other words (19) is an instance of *non-imposition*. Similarly in (20), by not renewing their subscriptions the members abstained from imposing a barrier to their resigning their membership. Even if our world knowledge does not include an acquaintance with the rules of horse-racing, the fact that the race starter has been subjected to criticism in (21) allows us to infer that he should have imposed a barrier to the horses' approaching too close to the starting tape. These three tokens do not require any further knowledge of the co-text in order to determine the type of permission involved. (22) is different in this respect. It is only by acquainting ourselves with the co-text that we can be sure that (22) does not imply a prior prohibition on the descent of "the blacks". In fact (22) merely states that the permitters did not themselves make any effort to seek them out.

Tokens such as (22) may appear at first sight to be ambiguous. However, this sort of ambiguity usually evaporates when one conducts a thorough examination of the co-text. Whenever such an investigation reveals no clue as to the previous existence of a barrier to the realisation of the situation encoded in the complement clause, the token in question is labelled as encoding *non-imposition*. The question of the presence or absence of a barrier is a black-and-white question. Either such a barrier existed, or it did not. If it existed one may expect it to have been either explicitly mentioned or at least implied by the speaker.

Examples (13)–(22) show that both *barrier-removal* and *non-imposition* may be encoded using both *allow* and *let*. How often are the two constructions actually used to encode the two sorts of permission? The answer is shown in Table 3, in which we see that while *allow* is used to encode *barrier-removal* in almost nine cases out of ten, *let* favours *non-imposition* by a margin of almost four to one.

Table 3. Constructions containing positive active voice matrix verbs *allow* and *let* encoding *barrier-removal* or *non-imposition* with horizontal percentages

	Totals per sample		**Percentage totals**	
Construction	***barrier-removal***	***non-imposition***	***barrier-removal***	***non-imposition***
allow to-inf.	365	49	88.2%	11.8%
let bare inf.	81	291	21.8%	78.2%

The difference between the two constructions with respect to encoding *barrier-removal* or *non-imposition* is statistically significant ($p < 0.0001$). Moreover, there is not as much overlap between the two constructions, especially in the encoding of *barrier-removal*, as appears at first sight in Table 3. Of 81 tokens of *let* encoding *barrier-removal*, as many as 25 contain the predicate *go*, while among the 365 tokens of *allow* encoding *barrier-removal*, on the other hand, just one contains the predicate *go* and this one does not encode the *release* sense, which is the most common meaning of '*let* x *go*'. In addition, of the remaining 56 examples of *barrier-removal* encoded by *let*, another 27 contain other motion verbs, such as *drop*, *slide*, *visit* and *come*. The prototypical sort of barrier in the case of *barrier-removal* readings of *let* is thus one prohibiting physical movement.

We turn now to active voice constructions containing negative polarity matrix verbs *let* and *allow*, which encode either *barrier-retention* or *barrier-imposition*. These two forms of (refusal of) permission are illustrated in Figures 3 and 4.

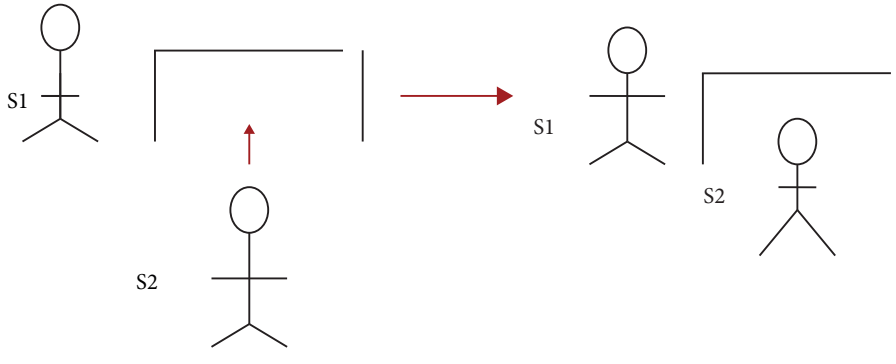

Figure 3. Retention of barrier by S1 hinders S2 from passing

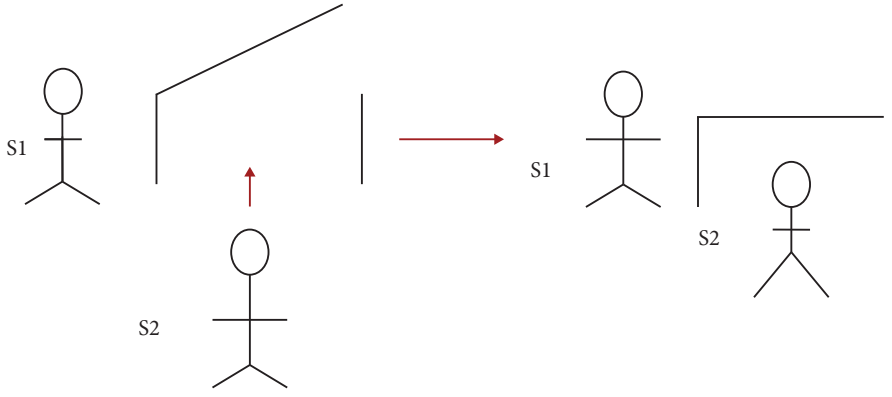

Figure 4. Imposition of barrier by S1 hinders S2 from passing

The criteria for distinguishing between *barrier-retention* and *barrier-imposition* are similar to those used to distinguish between *barrier-removal* and *imposition*. We again find that both types of prohibition may be encoded by both constructions as exemplified by the tokens of *barrier-retention* in (23) and (24) and *barrier-imposition* in (25) and (26).

(23) It is our interests, rather than those of a degenerate and selfish minority, that the police should protect; and if the law at present *does not allow them to do so* then the law must be changed. (BNC C88 1105)

(24) They *don't let women drive* cars, let alone fly an aircraft. (BNC BNV 987)

(25) '*Don't let her get away*, Tim!' he shouted. (BNC B0B 478)

(26) After the feud he *refused to allow Jamila to visit* her parents. (BNC A6V 790)

Table 4, which may be compared to Table 3, contains details of how often the two constructions are used to encode the two types of prohibition.

Table 4. Constructions containing negated active voice matrix verbs *allow* and *let* encoding *barrier-retention* or *imposition*

	Totals per sample		**Percentage totals**	
Matrix verb	***barrier-retention***	***imposition***	***barrier-retention***	***imposition***
allow	21	27	43.8%	56.3%
let	25	93	21.2%	78.8%

The totals in Table 4 indicate that there is a greater degree of overlap between the two constructions with negated matrix verbs than was the case with positive ones,

as shown in Table 3. Nevertheless, the difference between the two constructions with respect to encoding *barrier-retention* or *imposition* is still statistically significant ($p < 0.001$), indicating that the two are by no means always interchangeable. Taken together, Tables 3 and 4 provide eloquent testimony to there being a clear difference of meaning between the permissive constructions containing *let* and *allow*.

3. Constructions with passive matrix verbs

While active permissive *allow* and *let* are both very common, their passive counterparts differ greatly in this respect. Passive *allow*, as in (1), reproduced here for convenience, resembles active *allow* in so far as it is one of the half dozen most common (passive) matrix verbs in English. Passive *let*, on the other hand, as in (2), is extremely rare, being represented by only 19 relevant tokens (of 22 in all) in the whole of the BNC.[2]

(1) Few towns can boast they do not have a problem with dogs *being allowed to roam* the streets and Darlington is no exception. (BNC K54 6237)

(2) She *shouldn't be let roam* the hills alone. (BNC G0X 7)

(1) is an example of *non-imposition*, nothing having been done to stop the dogs from roaming. (2) is an example of *barrier-imposition*, the speaker expressing the opinion that a barrier ought to be implemented to prevent the subject's roaming. Just as in the case of active matrix verbs, we also find both *non-imposition* and *barrier-removal* encoded by passive *let*, as in (27) and (28). Eight of in all 11 tokens of barrier-removal contain the verb *go*.

(27) 'Things *were being let slide* because it was due to close in five weeks time.' (BNC K3K 401)

(28) Some relationships *have to be let go* in order that new ones can flourish. (BNC BNF 1571)

There are four tokens of *barrier-imposition* encoded by passive *let* in the BNC, one of which has been cited as (2), but none of *barrier-retention*.

All four forms of permission and prohibition are found encoded by passive *allow*. An instance of *non-imposition* has been cited as (1). *Barrier-removal* is exemplified in (29), *barrier-imposition* in (30) and *barrier-retention* in (31). The total numbers of tokens for both matrix verbs are given in Table 5.

2. The irrelevant tokens are the [MAKE REDUNDANT] sense of 'let go', which is causative rather than permissive, in so far as the person dismissed, the causee, has no say in the matter.

Table 5. The number of tokens of four types of permission/prohibition with passive *allow*and *let* in two samples with vertical percentages

	***'be let bare*-infinitive' in BNC**		***'be allowed to*-infinitive' in random sample of 1,000 tokens of *allow***	
Barrier-removal	11	58%	48	34%
Non-imposition	4	21%	41	29%
Barrier-retention	0		17	12%
Barrier-imposition	4	21%	35	25%
Total	19		141	

(29) For years Judaism was suppressed in the Soviet Union, practised behind closed doors, and often in fear. The school *was allowed to open* only nine months ago. (BNC KRU 225)

(30) She *hadn't been allowed to bring* anything off the boat except her patchwork leather shoulder-bag which had been thoroughly searched first. (BNC H7W 113)

(31) Because of the Sex Discrimination Act they're *not allowed to advertise* a women only service or recruit only women drivers. (BNC K26 1622)

An obvious question prompted by the data in Table 5 is why there are so many tokens of *barrier-removal* encoded by passive *let* when active *let* overwhelmingly favours *non-imposition*? The answer to both these questions lies, I suggest, in the semantics of *non-imposition*. Figures 5 and 6 illustrate passive *barrier-removal* and *non-imposition*, i.e. situations in which the permitter is not explicitly encoded.

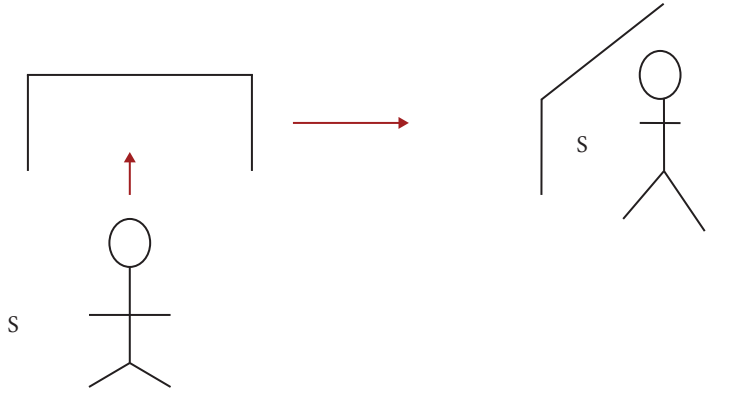

Figure 5. *Barrier-removal* enabling the permittee (S) to pass

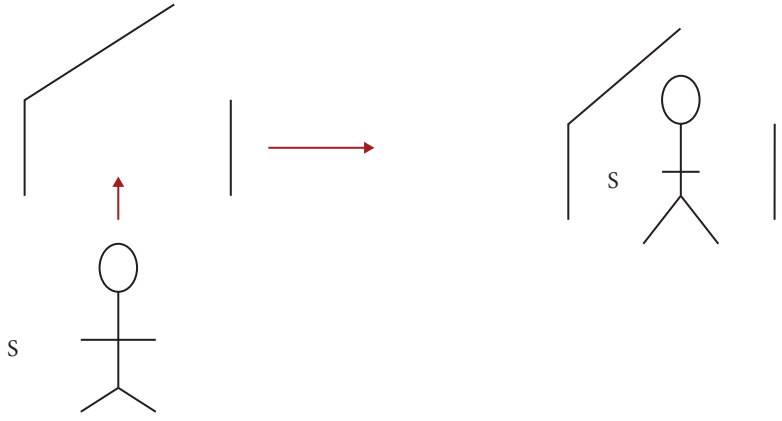

Figure 6. *Non-imposition* enabling the permittee (S) to pass

In Figure 5 a barrier is removed, enabling the permittee to move unhindered on his or her way. In Figure 6, on the other hand, a barrier is seen to remain unlowered. Moreover, there would appear to be little reason to encode the possibility of its being lowered. In Figure 2, which illustrates the situation with an active matrix verb, this possibility may be inferred from the very presence of the permitter. However, in situations such as the one illustrated in Figure 6 there is little motivation for explicit encoding of a possible (lowered) barrier.

This explanation, however, raises another question, which is why there are so many tokens of *non-imposition* encoded by *allow*. After all, if Figure 6 accurately reflects the situation pertaining to cases of *non-imposition*, should not such cases be equally rarely encoded by the '*allow to*-infinitive' construction? In fact, as shown by Table 5, this is not the case. The difference between *let* and *allow* may be ascribed, I think, not to the matrix verbs themselves, but to the form of the complement clause, in other words to the difference between the semantics of the bare and the *to*-infinitive. In Egan (2008: 99) a *to*-infinitive complement is said to encode "a situation, viewed as a whole [and] profiled as the more/most likely of two or more alternatives in some specified domain". In other words the *to*-infinitive always encodes a targeted alternative, with one or more alternative situations lurking in the background, as it were. Figure 7, which illustrates *non-imposition* encoded by a passive matrix verb and a *to*-infinitive complement, incorporates the element of a latent alternative.

Is there any evidence of the implication of such latent alternatives among the tokens of '*allow to*-infinitive' in the corpus?[3] Five of the 41 relevant tokens are in

3. The notion of latent alternative implies an element of choice on the part of a human permittee. In (33) the fact that one is allowed to use a jacket, does not imply that one is

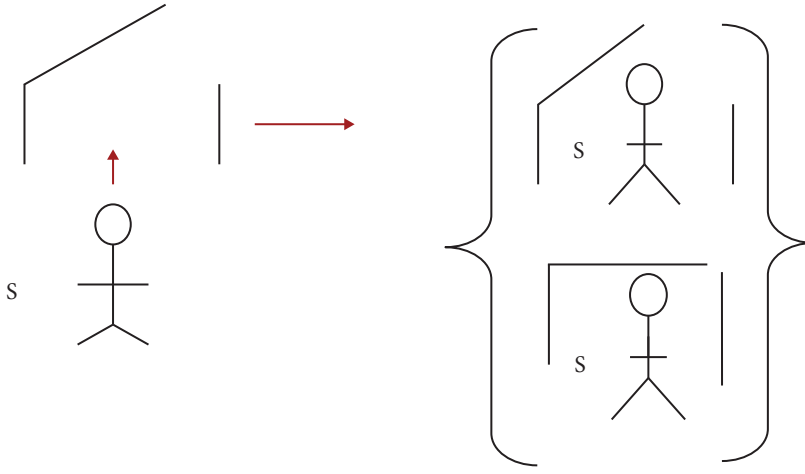

Figure 7. *Non-imposition* enabling the permittee (S) to pass, with an implied latent alternative of *barrier-imposition*

if-clauses, as in (32), and five in questions, as in (33). In these cases there is a clear implication of a latent alternative to the situation actually realised in the complement clause.

(32) If the teeth *are allowed to become sharp*, the cheek then becomes bruised and cut, causing pain thus making it difficult for the horse to chew properly. (BNC BPB 852)

(33) *Are* you *allowed to use* bulletproof jackets? (BNC FM7 942)

There are no *if*-clauses or questions among the 4 tokens of *non-imposition* '*be let* bare infinitive'. Among the 11 tokens of *barrier-removal* '*be let* bare infinitive' there is one *if*-clause. Among the 48 tokens of *barrier-removal* '*be allowed to*-infinitive', there are no *if*-clauses and only one question.

Some other examples of *non-imposition* encoded by passive *allow* are cited as examples (34)–(38). To what extent are latent alternatives implied in these tokens?

(34) Family Division President Sir Stephen Brown, making an open court statement after an hour-long private hearing, said: 'I do hope the child *will be allowed to continue* her life in these present circumstances in peace and without any form of harassment.' (BNC K5D 11288)

(35) I remain astonished that this state of affairs *is allowed to exist.* (BNC CH1 8165)

obliged to use one. In other words the construction is [+ Choice], to adopt the term used by Rudanko (2014).

(36) So far, Mr Berlusconi *has been allowed to keep* three national TV channels – the same number as RAI. (BNC CRC 2418)

(37) Expert witnesses *are* usually *allowed to remain* in court during the testimony of other experts in their field, and sometimes throughout the hearing if it is important that they hear all the evidence. (BNC J76 852)

In (34) the fact that the speaker expresses a hope that the complement situation may continue without interruption implies that there is a real possibility of this not happening. In (35) the fact that the continued existence of the complement situation arouses astonishment in the speaker implies that it should be brought to a halt. From the adverbials *so far* in (36) and *usually* in (37) we may infer that the realisation of the complement situation may be blocked in certain circumstances. In all four of these tokens the existence of a latent alternative is strongly implied by the speaker.

Turning our attention to negated passive matrix verbs, we saw in Table 5 that *barrier-imposition*, illustrated in Figure 8, may be encoded by both *let* as in (38) and *allow* as in (39).

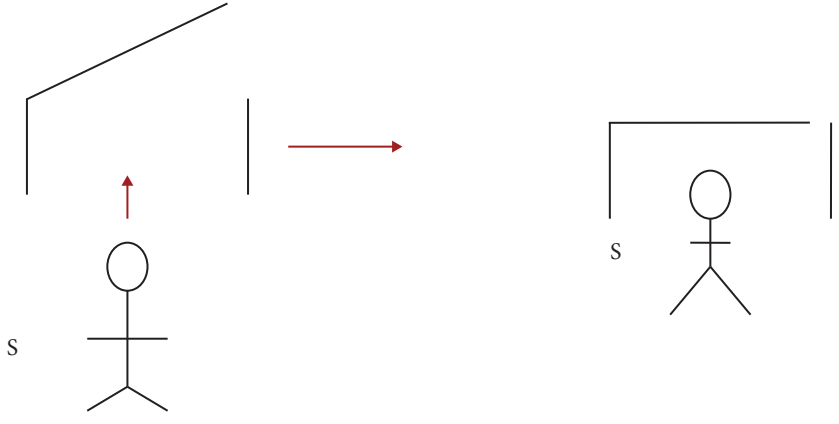

Figure 8. Imposition of barrier hinders S from passing

(38) When I was left at school I was savage at *not being let go home*; and when I went home, my mother did nothing but find fault with my schoolboy manners. (BNC HXG 917)

(39) As it turned out, Mario *wasn't allowed to race* at Monza because he'd driven a dirt-track race within the previous twenty-four hours, and it wasn't until Watkins Glen at the end of 1968 that he first drove in a FI race. (BNC CD9 1448)

Neither (38) nor (39) encode a permanent ban on home-coming or racing as evidenced by the adverbials *When I was left at school* in the former and *As it turned*

out in the latter. They are therefore classified as instances of *barrier-imposition* rather than *barrier-retention*. The latter form of prohibition with passive matrix verbs is illustrated in Figure 9.

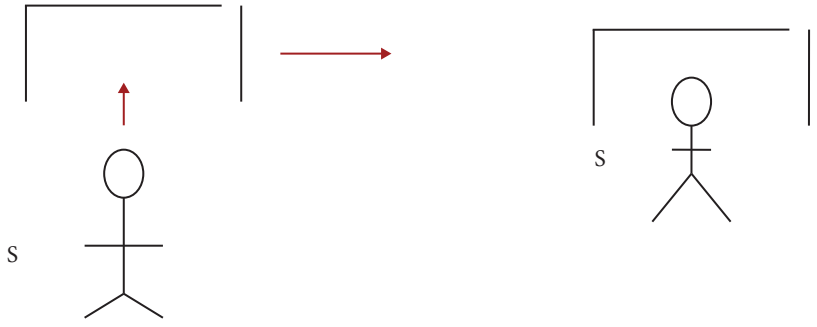

Figure 9. Retention of barrier hinders S from passing

The situation in Figure 9 is not encoded at all by *let* in the BNC, presumably for similar reasons to those adduced in the case of *non-imposition* above. Basically there is very little happening for the speaker to encode. In the case of the passive *allow* construction, on the other hand, the *to*-infinitive form of complement implies the possibility of a latent alternative, as illustrated in Figure 10, which may be compared to Figure 7.

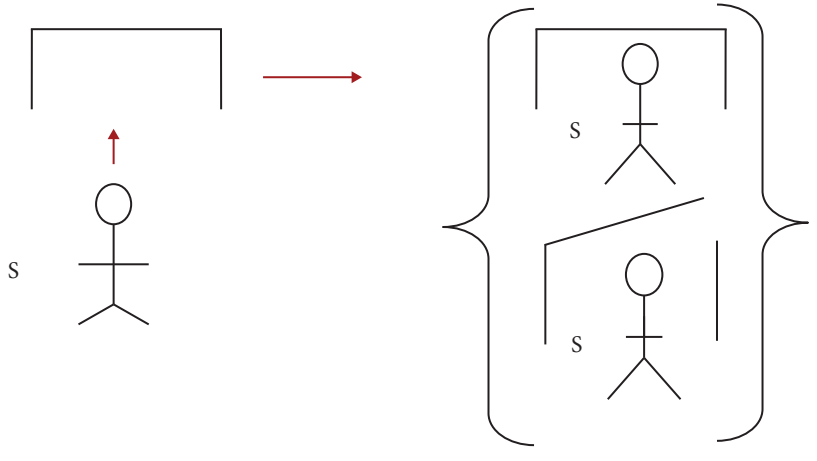

Figure 10. *Barrier-retention* hindering the permittee (S) from passing, with an implied latent alternative of *barrier-removal*

One example of *barrier-retention* encoded by '*be allowed to*' has already been cited as (31). Other typical examples are (40)–(43).

(40) During that time *no* Chadian resident *was allowed to seek* information about the prisoners, as they risked becoming prisoners themselves. (BNC CJP 23)

(41) The press *are not* normally *allowed to be present* during chambers applications. (BNC J76 824)

(42) The Club will make the necessary arrangements, but *no-one is allowed to go* into town before clearing immigration nor should the skipper or any of the crew visit immigration as they will be told, in no uncertain terms, to return to the club immediately. (BNC G37 606)

(43) Magistrates who deal with family matters have been specially trained and *are not allowed to sit* in the Family Court until that training has been completed. (BNC B03 1986)

All four tokens (40)–(43) encode situations in which barriers have not been raised, specifically in order to hinder the permittee from proceeding. They all, however, also contain adverbials (underlined) which imply that there may exist circumstances in which the barrier in question might be raised. Thus in (40) the adverbial *During that time* allows us to infer that the barrier to seeking information may have been lifted at a later date. In (41) the adverbial *normally* allows us to infer that the barrier to the presence of the press may be lifted in exceptional circumstances. In (42) the adverbial *before clearing immigration* allows us to infer that the barrier will be removed when this proviso is satisfied. Similarly, the adverbial *until that training has been completed* in (43) allows us to infer the future possibility of the rescindment of the prohibition. Seven out of a total of 17 tokens of *barrier-retention* encoded by passive *allow* contain this sort of adverbial as opposed to just three of 35 tokens of *barrier-imposition* passive *allow* and none of the four tokens of *barrier-imposition* passive *let*. There are, as we have already seen, no tokens of *barrier-retention* encoded by passive *let*. A Fisher's Exact text shows that the difference between the two sorts of prohibition in this respect is statistically significant at the level of $p = 0.01$.

4. Passive *let* in COCA

In Section 3 we saw that the passive *let* construction is used in the BNC for both forms of permission and one of two forms of prohibition. In this section we ask whether American English, as represented in COCA, displays the same distribution. The numbers for COCA are given in Table 6, with the numbers for the BNC repeated from Table 5 for ease of comparison.

Table 6. The number of tokens of four types of permission/prohibition with passive *let* in the BNC and COCA

	'*be let* bare infinitive' in COCA		**'*be let* bare infinitive' in BNC**	
Barrier-removal	79	81%	11	58%
Non-imposition	15	15%	4	21%
Barrier-retention	1	1%	0	
Barrier-imposition	3	3%	4	21%
Total	98		19	

Since there are at time of writing some five times as many words in COCA as the BNC (530m. compared to 100m.), we can conclude that the construction is equally common in the two varieties. American English appears to differ from British English in employing the '*be let* bare infinitive' construction to code *barrier-retention*, if one can draw this conclusion on the basis of a single example, cited here as (44), which encodes the continuation of imprisonment, i.e. the absence of [RELEASE]. As many as 68 of tokens of *barrier-removal* also contain '*be let go*' coding the [RELEASE] sense, exemplified by (45) and (46).[4]

(44) He'd *never be let go*, but Sigmund acted as though he believed. (COHA Fiction 2008)

(45) After each snake had been marked, it *was let go.* (COCA Fiction 2009)

(46) She *was let go* after being questioned by police. (COCA News 2001)

Not all tokens of '*be let go*' permissives encode the [RELEASE] sense. There are tokens of *barrier-removal*, which do not involve releasing, as in (47), as well as tokens coding *non-imposition* as in (48), and *barrier-imposition* as in (49).

(47) I actually felt as though this nuisance that was hanging onto me for these two years was now *being let go* and I could move forward with my life. (COCA Spoken 1991)

(48) The house we moved into *had been let go*. (COCA Fiction 1998)

(49) And that's another focus of this thing that *should not be let go*. (COCA Spoken 1994)

(47) encodes the removal of a barrier to the speaker's having surgery to amputate an arm and shoulder, which doctors for several years had tried to save. (48) encodes

4. In addition there are 355 passive tokens of the causative [MAKE REDUNDANT] sense of '*let go*'.

the deterioration of the property when steps had not been taken to stop this, in other words *non-imposition*, and (49) the need to put in place a mechanism to stop something (Bill Clinton's peccadillos) being forgotten, i.e. *barrier-imposition*.

There are only seven tokens of '*be let* bare infinitive' that contain verbs other than *go*, and four of these contain *stay*. Two of these encode *non-imposition*, as in (50), and two *barrier-imposition,* as in (51).

(50) […] this was so even when you were practically sure you *would be let stay* on for another ride. (COCA Fiction 2007)

(51) If he'd known then what he does now, Dawson Kalliam wouldn't have been exiled and Feldin Maas *wouldn't have been let stay*. (COCA Fiction 2012)

The final three tokens contain the verbs *run*, *die* and *be*, all three of which encode *non-imposition*.

We turn now to the passive *let* construction containing the *to*-infinitive rather than its bare counterpart, exemplified here by (52)–(54).

(52) I'm sure they were stopped and questioned on their way out, but my understanding is they *were let to leave* the property. (COCA Spoken 1966)

(53) Sir, we enlisted men can't resign. That'd be desertion. But the officers *are let to walk off* whenever they like. (COCA Fiction 2003)

(54) My daddy Strother didn't credit it, though, and he beat Mama near about to death, saying nothing that piddling could be his get, allowing as how I maybe wasn't even human and *should not be let to live*. (COCA Fiction 2003)

(52) is an example of *barrier-removal*, (53) of *non-imposition*, nothing being done to stop the officers leaving, and (54) of *barrier-imposition*, the father wishing to terminate his son's life.

According to the *OED* "A few examples of the use of *to* before the infinitive in this construction occur in all periods; now chiefly when *let* is used in the passive" (definition 12, b.II.). As pointed out in the Introduction, the only token of this construction in the BNC actually occurs in a quotation from a nineteenth century text. There are, however, twelve tokens like (52)–(54) from Present-day English in COCA. Details are given in Table 7, with numbers for '*be let* bare infinitive', repeated from Table 6.

We have seen that a large majority of tokens of '*be let* bare infinitive' instantiate the (at least partially) lexicalised [RELEASE] sense of '*let go*'. This sense is not found at all with '*be let to*-infinitive'. It may well be pre-empted by the entrenched status of the construction with the bare infinitive. If we exclude tokens coding the [RELEASE] sense in an effort to compare more like with like, we arrive at the numbers in Table 8.

Table 7. The number of tokens of four types of permission/prohibition with passive *let* bare infinitive and passive *let to*-infinitive in COCA

	'*be let to*-infinitive'		'*be let* bare infinitive'	
Barrier-removal	5	43%	79	81%
Non-imposition	4	33%	15	15%
Barrier-retention	0		1	1%
Barrier-imposition	3	25%	3	3%
Total	12		98	

Table 8. The number of tokens of four types of permission/prohibition with 'passive *let* bare infinitive', minus the [release] sense of '*let go*' and '*let to*-infinitive' in COCA

	'*be let to*-infinitive'		'*be let* bare infinitive'	
Barrier-removal	5	43%	11	38%
Non-imposition	4	33%	15	52%
Barrier-retention	0		0	
Barrier-imposition	3	25%	3	10%
Total	12		29	

The distribution of the two constructions shown in Table 8 does not involve any statistical difference. There is, however, a difference between the two, which is not apparent from the table. This is the type token ratio. Whereas the 29 tokens with the bare infinitive span over just five verbs, with *go* and *stay* accounting for all but three tokens, the 12 tokens containing the *to*-infinitive instantiate 11 different verbs. These include change of location and change of state verbs like *vanish*, *leave*, *pass* and *die* (the only verb to occur twice), but also stative verbs like *abide* and *live*. In this respect they resemble the '*be allowed to*-infinitive' construction rather than the '*be let* bare infinitive' construction.

So far all the examples but one of the two passive *let* constructions have been taken from fictional or spoken texts. One may wonder whether these texts are typical for the constructions. Details of the text types in which both occur are given in Table 9. Note that the News and Magazine categories have been merged in the table and that tokens labelled News which occur in interviews have been assigned to the Spoken category.

We see in Table 9 that passive *let* constructions may be found in many genres. Two of the academic examples, cited as (55) and (56), are from the *Anthropological Quarterly*.

Table 9. The number of tokens of 'passive *let* bare infinitive' and '*let to*-infinitive' according to the various text types in COCA

	'*be let to*-infinitive'		**'*be let* bare infinitive'**	
Spoken	3	25%	44	45%
Fiction	7	58%	19	19%
News/Magazine	2	17%	32	33%
Academic	0		3	3%
Total	12		98	

(55) The water *must be let run* from the crotch of the husband's trousers down to the shoe. (COCA Academic 2005)

(56) According to many Thai Buddhists […], love (rak) is a kind of attachment. It can bind us unhealthily to the object or person loved, and for that reason it *may need to be let go of*. (COCA Academic 2005)

These two examples are from different issues of the journal, (55) from an article entitled "Rethinking the Couvade", and (56) from an article entitled "Orthodox Hybridities: Anti-Syncretism and Localization in the Evangelical Christianity of Thailand". The reason for citing the titles is to give some impression of the register employed in the articles, which is obviously far from the norm of the spoken language. Having said that, both tokens appear completely idiomatic, at least to the ears of the present writer. Both passive *let* constructions, although they may be rare, would seem to be perfectly acceptable in a variety of registers in Present-day English. In the next section I will trace their evolution in American English over the past two centuries.

5. Passive *let* in COHA

The numbers of tokens of both passive *let* constructions in COHA are given in Table 10. The data have been divided between three periods with approximately the same number of words.

According to Table 10, both constructions were almost twice as frequent (in terms of normalised frequencies) in the nineteenth than the twentieth century. This difference may be ascribed to the total incidence of tokens coding *barrier-removal* in the two periods. If we subtract these from the totals, there is no appreciable difference between the three periods in the table (the probability of their coming from similar populations is over .75 according to a Fisher Exact Test). (57)–(59) are typical examples of the *barrier-removal* sense in the earliest period.

Table 10. Both passive *let* constructions in COHA

	Pre-1901		1901–1960		Post-1960	
	Approx. 130m. words		Approx. 140m. words		Approx. 130m. words	
	be let V	*be let to V*	*be let V*	*be let to V*	*be let V*	*be let to V*
Barrier-removal	44	5	28	1	16	0
Non-imposition	12	9	11	5	9	4
Barrier-retention	0	0	0	1	1	0
Barrier- imposition	9	2	4	2	4	2
	65	16	43	9	30	6

(57) Just as the anchor *was let go* a signal gun was fired. (COHA Memoirs 1837)

(58) […] then, as her head pointed quartering down the river, the stern line *was let go*, and we shot away like an arrow from a bow. (COHA Magazine 1879)

(59) The canvas was carried clean from the bolt-ropes, the sheets *were let go*, and the lighter sails clewed up. (COHA Fiction 1868)

The majority of tokens coding *barrier-removal* in the nineteenth century contain the verb *go* used in a nautical context. Many of these usages would be rendered redundant with the decline of sailing ships.

Unlike the case of the '*be let* bare infinitive' construction, there are very few instances of *barrier-removal* in COHA encoded by the '*be let to*-infinitive' construction, especially in the twentieth century. As for *non-imposition*, it is found in both constructions, with approximately twice as many examples containing the bare infinitive. (60) and (61) are two nineteenth century examples containing the same verb, *grow*.

(60) Every moment this hatred *is let grow* in the heart's garden, it spreads and strengthens, till it gains dominion and makes men slaves, and madder than before. (COCA Fiction 1874)

(61) It was her one, eager, passionate longing, in these childish days, that these locks of hers *should be let to grow*. (COCA Fiction 1863)

Grow is one of only three verbs to occur with both forms of complement. The others are *go* and *live*. We should also note that, unlike the case in COCA, discussed in Section 4, the type token ratio is similar for the two constructions. If one leaves aside the verb *go*, which is very common in the bare infinitive construction, with 112 tokens, and much less common in the *to*-infinitive construction, with just seven examples, the most recent of which is from 1905, there are then 26 tokens

of '*be let* bare infinitive', containing 22 different verbs, and 24 tokens of '*be let to*-infinitive', containing 19 different verbs.[5] Both constructions are used with motion verbs, such as *run*, *recede* and *come* in the bare infinitive construction, and *walk* and *ride* in the *to*-infinitive construction. However, they are also used with many other sorts of verbs. Moreover both constructions may be used to encode all four of Vendler's (1967) situations types, illustrated here by (62)–(65), all containing the *to*-infinitive construction.

(62) [Women] see their blood, and it does them good, while men *are let to be* vainer. (COHA Fiction 1953)

(63) I'm sorrier than I can tell that ever you *were let to fool* with powder. (COHA Fiction 1900)

(64) […] so the earth *was let to bring forth* animals in the living likeness of itself. (COHA Fiction 1966)

(65) They have certain legends that must be preserved for their public and truth so much more fascinating than fiction in most of their cases *must be let to drop* by the wayside. (COHA Magazine 1928)

(62) encodes a stative predication and (63) an activity, (64) an accomplishment and (65) an achievement.

Before rounding off this section, mention must be made of two tokens that encode *barrier-retention*, a form of prohibition that does not occur with passive *let* in the BNC. One of these, (44), is also included in COCA and is repeated here for convenience. The other, cited as (66), contains the *to*-infinitive.

(44) He'd *never be let go*, but Sigmund acted as though he believed. (COHA Fiction 2008)

(66) But there's times, Ann, when just for a bit they're just like children. They need comforting without *being let to know* they are being comforted. (COHA Fiction 1913)

In (44) the subject is being held captive, and states his belief that the prohibition on his being released will be maintained. As for (66) I stated in Section 2 that tokens of the '*let* x *know*' (= 'inform x') construction were omitted from this study. However, (66) is not an instantiation of this multi-word verb. Rather *know* here means

5. One may note in passing that COHA contains 64 tokens of the passive causative '*be let go*' construction, meaning [MAKE REDUNDANT]. This construction seems to have been first used between the world wars, with all but four examples occurring after 1950.

[REALISE]. The meaning of (66) is that the permittees (men!) should be allowed to remain in ignorance of the fact that they are receiving comfort: in other words that the veil disguising this fact should not be lifted.

6. Summary and conclusion

In this paper I have addressed the question of why there are so few *let* passives and have suggested that the answer is related to the fact that *let* prototypically encodes the form of permission which I have termed *non-imposition*. When the matrix verb is in the active voice the situation encoded is construed as one in which the permitter refrains from acting, thus allowing the complement situation to evolve: in other words '*x* did nothing to stop *y* happening'. When the matrix verb is in the passive, however, *x* is airbrushed from the picture, so to speak, leaving us with 'nothing occurred to stop *y* happening'. Given that this statement could be applied to all situations in which something occurs, it cannot be said to be very informative. Hence it tends to be avoided. This avoidance is not, however, total. We have seen in the data from the BNC and COCA that both types of permission (*barrier-removal* and *non-imposition*) are encoded, albeit not frequently, by the '*be let* bare infinitive' construction. Of the two types of prohibition (*barrier-retention* and *barrier-imposition*), both can be found in COCA, but only *barrier-imposition* in the BNC. We have also seen that all four types are encoded in the COHA data by the even rarer '*be let to*-infinitive' construction.

I have argued that the difference between the passive constructions with *allow* and *let* in the BNC may be ascribed to a difference in the form of the infinitive complement. Whereas the bare infinitive merely encodes a situation as a whole (as described by Langacker 1990: 82), its *to*-infinitive counterpart encodes the targeted of several possible alternatives. It is the presence of latent alternatives in the background, as it were, that licences the use of the passive *allow* construction to encode the relatively content-less situation of permitter-free *non-imposition*. Similarly, it is the presence in the background of the alternative of *barrier-removal* that licences the use of passive *allow* to encode *barrier-retention*, while one of only two examples in COHA of passive *let* used to encode this form of prohibition also contains the *to*-infinitive.

Acknowledgements

I would like to thank the editors of this volume for inviting me to submit a chapter in honour of Juhani Rudanko, a pioneering scholar in the field of complementation studies. An earlier version of Sections 2 & 3 was published as Egan (2009). Thanks are due Marc Fryd, Jérôme Puckica

and Jean Claude Souesme for their comments on this earlier version, and to Lalita Murty and two anonymous reviewers for their comments and advice on the present one.

References

Primary sources

The British National Corpus. 2001. Oxford: Oxford University Computing Services.
The Corpus of Contemporary American English. <http://corpus.byu.edu/coca/> (June 2016).
The Corpus of Historical American English. <http://corpus.byu.edu/coha/> (June 2016).

Secondary sources

Carter, Ronald & McCarthy, Michael. 2006. *Cambridge Grammar of English: A Comprehensive Guide: Spoken and Written English Grammar and Use*. Cambridge: CUP.

Egan, Thomas. 2003. *Distance and Direction: A Usage-based Study of Infinitive and -ing Complement Clauses in English* [Acta Humaniora nr. 170]. Oslo: Unipub Forlag.

Egan, Thomas. 2008. *Non-finite Complementation: A Usage-based Study of Infinitive and -ing Clauses in English*. Amsterdam: Rodopi. doi:10.1163/9789401205542

Egan, Thomas. 2009. Why are there so few *let* passives? In *The Passive in Germanic Languages*, Marc Fryd (ed.). Special issue of GAGL: Groninger Arbeiten zur Germanistischen Linguistik. <http://gagl.eldoc.ub.rug.nl/root/2009-49/2009-49-12/>

Kemmer, Suzanne & Verhagen, Arie. 1994. The grammar of causatives and the conceptual structure of events. *Cognitive Linguistics* 5(1): 115–156. doi:10.1515/cogl.1994.5.2.115

Langacker, Ronald. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.

Oxford English Dictionary, 2009, 2nd edn [CD-Rom]. Oxford: OUP.

Rudanko, Juhani. 2014. A new angle on infinitival and *of -ing* complements of *afraid*, with evidence from the *TIME* Corpus. In *Corpus Interrogation and Grammatical Patterns* [Studies in Corpus Linguistics 63], Kristin Davidse, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds), 223–238. Amsterdam: John Benjamins.

Talmy, Leonard. 1986. Force dynamics as a generalization over 'causative'. In *Languages and Linguistics: The Interdependence of Theory, Data, and Application*, Deborah Tannen & James E. Alatis (eds), 68–75. Washington DC: Georgetown University Press.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Cambridge MA: The MIT Press.

Thomson, Audrey Jean & Martinet, Agnes V. 1986. *A Practical English Grammar*, 4th edn. Oxford: OUP.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca NY: Cornell University Press.

Goldberg's Rely On construction

Overreliance on generalization?

Jouni Rostila
University of Helsinki

Goldberg (2014) proposes one polysemous argument structure construction (ASC) for cases as varied as *nibble/rely/bet on something*. Inspired by ASCs needed for a semantically similar domain in German, my analysis suggests that a more adequate solution can be reached with three constructions that are semantically further apart than the two sub-senses of Goldberg's Rely On construction. The solution makes use of Israel's (1996) empirical findings regarding the historical development of the English *way* ASC to model the development and interrelationships of the required constructions. Overall, the chapter advocates the advantages of a contrastive approach and the use of diachronic studies to inform synchronic Construction Grammar analyses of specific domains.

Keywords: Construction Grammar; argument structure construction; preposition; contrastive linguistics; German; English

1. Introduction

A crucial question in Construction Grammar (henceforth CxG; cf. e.g. Hilpert 2014 for basics) analyses is how to split areas of grammar into form-meaning pairs, i.e. constructions. The same strings can often be generated by combining lexical constructions with either a few semantically general schematic constructions or with a number of more specific ones. While the former approach is attractive in that it may lead to elegant solutions capturing the broadest generalizations, the latter might be psychologically more valid and avoid overgeneration (cf. Croft 2001: 5; Boas 2003, 2011; Perek 2015: 214). Since Goldberg's (1995) introduction of argument structure constructions (ASCs), meaningful complementation patterns into which individual verbs are embedded (cf. Rostila 2015: 34–36), the choice between these options within the realm of verb complementation has been subject to a lively debate, cf. e.g. Müller and Wechsler (2014) and the responses to this target article such as Boas (2014). It

DOI 10.1075/slcs.195.04ros

seems that ASCs are sometimes semantically too general and thus overgenerate, cf. e.g. Boas' (2003, 2011) criticism of Goldberg's (1995) resultative construction for English. However, this is no reason to reject ASCs categorically – only a reason to be careful to find the adequate level of generalization for the patterns to be described.[1]

Goldberg (2014) proposes for English an ASC she calls the Rely On construction, which generates cases as widely varied as *nibble/gnaw/feast on carrots, live on potatoes, prey on foreigners, chew on an idea, rely/depend on help from others, call on somebody to do something, bet on something*.[2] At least at first sight, the semantic spectrum covered by this construction – from verbs of ingestion (*nibble, gnaw*, etc.) via verbs of reliance to verbs with 'hope', i.e. prospective, semantics (*bet*) – might seem too broad for a single construction. In an attempt to find out whether this is the case, this chapter takes a closer look at Goldberg's proposal and contrasts it with ASCs proposed in Rostila (2007, 2014, 2015, in press) for similar areas of German verb complementation.[3] Moreover, the findings of Israel (1996) regarding the diachrony of the English *way* ASC are used as a basis for an alternative solution. The ultimate goal is by no means to impose German ASCs on English – such an undertaking would be futile, since each language is likely to have constructions of its own, cf. Croft (2001: 6) and Goldberg (2003: 222)[4] – nor to prove Goldberg's (2014) approach definitely wrong. Rather, the chapter aims at providing the basis for an alternative, improved account in terms of more specific ASCs, and in doing so brings to light more general problems of choosing between CxG analyses.

---

1. Cf. Boas (2014: 96), who stresses that generalizations – also ASCs – are needed on many levels; it is important just to find the right level.

2. Goldberg's (2014) proposal is couched in a comment article targeting Müller & Wechsler's (2014) arguments for a lexical approach to argument structure. Although the present chapter proposes an account based on more specific constructions than Goldberg's proposal, it does not take a stance against Goldberg's main arguments, but emphasizes the need to find the right level of specificity for ASCs. For some comments on the need to assume ASCs (contra Müller & Wechsler 2014), see Rostila (in press: n. 3).

3. See Boas (2010) for a broader view on contrastive CxG studies.

4. Shared history – shared parent language or language contact – may of course cause languages to have very similar constructions, but such constructions are hardly identical, since they have to function as parts of a different system. Both options are likely to be excluded in the case of the Rely On constructions, though: the German constructions considered probably only developed from the 17th century onwards (see Rostila 2016: 272), excluding common Germanic origin; also a contact scenario between English and German strong enough to cause the borrowing of an argument marking pattern is highly unlikely.

2. Goldberg's Rely On construction

A closer look at Goldberg (2014) makes it clear that it is not simply a matter of proposing one too broad a generalization that is at stake here. Rather, Goldberg (2014: 128) proposes two constructions joined by a polysemy link.[5] The essential properties of the "prototypical Rely On construction", or the "central sense" of the construction (ibid.), can be captured in the following way:

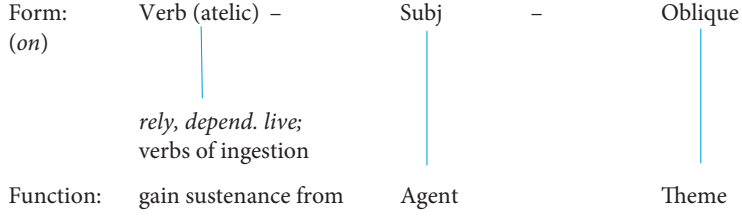

Figure 1. The Rely On construction: central sense

The figure, slightly adapted from Goldberg (2014: 128) to make it more easily comparable with those proposed here for German ASCs (see Section 3 below), is meant to express that the form of the construction consists of three "slots", one for a verb of a particular semantic type, one for a subject that has the function of Agent, and one for an oblique object marked by *on* functioning as Theme. Furthermore, the figure indicates that a verb embedded in the construction has the function of expressing the meaning 'gain sustenance from'; like ASCs in general, the construction can even impose this meaning on verbs not possessing it, cf. Goldberg (1995: 159; 2014: 127) and Rostila (in press: Section 2.1) for details. To gain a more concrete idea of the construction, it is useful to have a look at all the examples with which Goldberg (2014) illustrates the central sense:

(1) a. She nibbled on the roll.
b. The cow grazed/gnawed/chewed/dined/feasted/ munched/fed on apples.
c. She lived on potato chips/sushi/grass.
d. She lived on $10 a month.
e. The hyenas preyed on giraffes.
f. The landlord preyed on foreigners.
g. She chewed on the idea.

5. For different types of links between constructions, see Goldberg (1995: 72–81) and Hilpert (2014: 60–65).

The cases in (1) show that verbs of ingestion dominate in the examples, and – somewhat surprisingly – not a single example illustrates *rely,* the verb that is supposed to occur in the construction "quite frequently" (Goldberg 2014: 126).[6] While this bias in the examples does not necessarily lead to a decisively flawed account, one is still reminded of Rudanko's (1989: 148f.) criticism of an "impressionistic air" in linguistic work and his urge to strive for at least representative lists of verbs exemplifying a construction (ibid.). Section 3 will indeed show that considering more closely verbs of the *rely* type might have led to a different analysis.[7]

The following figure presents the essential properties of the second construction Goldberg posits, i.e. the "extended sense" of the Rely On construction (cf. Goldberg 2014: 128):

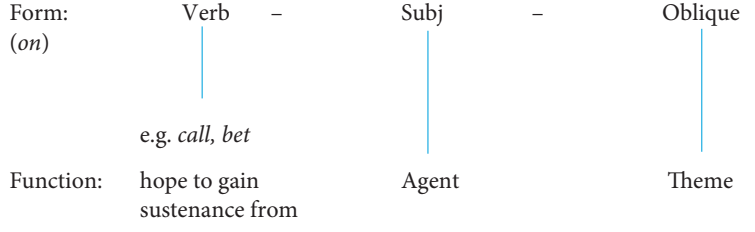

Figure 2. The Rely On construction: extended sense

Again, I have slightly adapted the notation from that used by Goldberg. Apart from the 'hope' semantics of the verb slot and the requirement of the central sense

6. The cases (1f)–(g) illustrate a metaphorical use of the central sense. The fact that precisely verbs of ingestion are semantically extended in this way might be an (admittedly weak) indication that ingestion constitutes the prototypical, and hence original, sense of the construction, cf. Section 4.

7. However, the present chapter certainly cannot solve the problem of compiling a representative list of verbs exemplifying the Rely On, and related, constructions; this remains a challenge due to the need to search for them on semantic grounds. Even FrameNet, a resource aiming at fairly comprehensive coverage of verb complementation on a semantic basis, only lists *count, depend* and *rely* as verbs evoking the Reliance Frame. Rudanko (1989: Ch. 6) yields a few further examples that stem from his analysis of Visser's (1973) type *I depended on him to come*: *call, count, depend, prevail, rely (on X to do Y); urge ((up)on X to do Y)*. Significantly, his analysis also suggests that there might exist a competing pattern with essentially the same semantics, but symbolized by the preposition *to*, cf. *John trusted to Mr. Smith, a lawyer, to draw up his will* (Rudanko 1989: 144).

that the verb be atelic, the extended sense inherits its properties from the central sense. Significantly, only *bet* and *call* are mentioned as examples of verbs occurring in the construction.

3. Preliminaries for an alternative solution

In my work on German ASCs (cf. e.g. Rostila 2006, 2007, 2014, 2015, in press), I have proposed ASCs based on prepositions of prepositional objects[8] that closely resemble the two senses of Goldberg's Rely On construction. First, German seems to display an ASC signified by the preposition *an* (+ dative) – interestingly, a historical cognate of English *on* (cf. OED,[9] *s.v. on*) – that similarly to the central sense of the Rely On construction combines with atelic verbs, e.g. verbs of ingestion:

(2) a. Er baute an einem Haus.
he built at a house
'He was building a house.'

b. Er schrieb/lies an einem Buch.
he wrote/read at a book
'He was writing/reading a book.'

c. Er trank an einem Bier.
he drank at a beer
'He was drinking a beer.'

d. Der Hund kaute an einem Knochen.
the dog gnawed at a bone
'The dog was gnawing a bone.'/'The dog gnawed on a bone.'

However, this construction differs in significant ways from Goldberg's central sense of the Rely On construction. First, it has a different range of application, combining not only with verbs of ingestion, but with other atelic verbs as well; on the other hand, the construction cannot express gaining sustenance, i.e. it does not occur with German translation equivalents of *rely, live,* etc. These make use of other prepositional object and/or lexical case structures instead, cf. *Er vertraut **seinen Freunden*** (dative)/***auf seinen Freunden*** (*auf* + dative), roughly 'He relies

8. Rostila (2007: Part II, Chapter 4; 2015: n. 9; in press: Section 1) show that it is necessary to differentiate between Ps of prepositional objects that correspond to lexical cases, Ps of this type that have grammaticalized into ASCs, and full local lexical Ps like *(be/appear)* ***on/in/at/over/under*** … (cf. Goldberg 2014: n. 7). The latter are fundamentally different and hence should not be put on a par with the former two categories.

9. *Oxford English Dictionary* online; <http://www.oed.com.06.08.2016>

on his friends'; *Er lebt* ***von*** *10 Euro am Tag* (*von* + dative) 'He lives on 10 Euros a day'. Second, the construction is used to express gradual progress, or progressive aspect, in an activity – hence I have labelled it the aspectual/incremental *an* construction (cf. Rostila 2006, 2007: 192f., 2015: 41).[10] In fact, it is probably this incremental/progressive semantics[11] that excludes verbs of sustenance from the construction: there can be no gradual progress in sustenance.

Given the differences, it is clear that proposing for English a construction semantically identical to the German ASC, but with the same formal pole as Goldberg's central sense, is no improvement over Goldberg's solution. However, the contrast to the similar German construction gives rise to useful questions about Goldberg's central sense. First, does it display similar progressive/incremental semantics? Cases like *The cow chewed on the apple for an hour – She nibbled on the roll* (Goldberg 2014: 126f.) suggest this. Second, is sustenance always a part of the semantics of cases that Goldberg considers examples of the central sense? It seems to me some of her examples, e.g. *chew, gnaw* and *nibble*, focus on incrementality/progressivity and at least background sustenance, if they do not exclude it altogether. On the other hand, when used with more abstract verbs, the *on* pattern seems to focus on sustenance and to background, or even exclude, progressivity/incrementality, cf. *feed, feast, dine, graze* and (1c)–(f).

If this semantic analysis is correct, the question arises whether it is justified to assume one construction for all these cases. A more viable alternative might be to assume one construction similar to the German incremental *an* for the cases focusing on progressivity/incrementality – only semantically narrower than in German, to account for the narrower range, i.e. the restriction to verbs of ingestion – and another one for the cases focusing on sustenance. Further still, given that there seems to be a semantic continuum between these two poles, it might be fruitful to assume a polysemy link between the two constructions. This aspect will be elaborated on in Section 4 on the basis of Israel's (1996) account of the development of polysemy in the English *way* ASC.

As regards the extended sense of Goldberg's Rely On construction ('hoping to gain sustenance'), it also resembles an ASC proposed here for German. As Rostila (2007, 2014, 2015, in press) show in more detail, there seem to be grounds for assuming an ASC in present-day German based on the prepositional object

---

10. Cf. Schøsler (2007) for a similar construction in Danish that interestingly is based on *på*, the Danish translation equivalent of *on.*

11. The division of labor between the aspectual *an* construction and the colloquial/dialectal German progressive (see Van Pottelberge 2004) is a relevant but complex issue that cannot be broached here.

preposition *auf* (+ accusative) that expresses prospectivity. The pattern occurs with prospective verbs like *warten* 'wait' and *hoffen* 'hope', cf. (3a); however, crucial proof of the existence of an ASC consists in cases like (3b)–(c), where the preposition *auf* can be seen to coerce a verb into a prospective meaning. A further indication of the semantic similarity of the German ASC to Goldberg's extended sense is that the pattern also occurs with the German meaning equivalent of the verb *bet*, which Goldberg (2014) gives as one of two examples of the extended sense, cf. (3d).[12]

(3) a. Er wartet/hofft auf einen Börsensturz.
he waits/hopes on a stock market crash
'He waits/hopes/is waiting/hoping for a crash of the stock market.'

b. Ich freue mich über/auf das Ende des Semesters.
I delight myself over/on the end of term
'I am glad of the ending of term' / 'I look forward to the end of term.'

c. Ich fahre auf Sieg, ganz klar.[13]
I drive on victory, quite clear
'I drive to win, that's clear.'

d. Er wettet auf Pferderennen.
he bets on horseraces
'He bets on horse races.'

Given the semantic similarity of this German ASC with Goldberg's extended sense, as well as the fact that in present-day German *auf* equals *on* in concrete local uses,[14] it seems tempting to assume a similar prospective ASC in English with the surface realization *on*. Section 4 puts this assumption on a more concrete footing by presenting further possible examples of such a pattern in English. It also makes an attempt to define the relationships of the three ASCs to each other that the analysis has hitherto suggested for Goldberg's Rely On cases. This section closes by showing the two German ASCs exploited in the analysis in more detail, cf. Figures 3 and 4:

---

12. See Rostila (2007, 2014, 2015, in press) for numerous further examples of *auf* + accusative with prospective predicates.

13. <http://www.bild.de/sport/motorsport/nico-rosberg/wehrlein-hat-uns-in-der-formel-1-sehr-geholfen-43060804.bild.html> (24 May 2016).

14. It must be emphasized that the correspondence of the concrete local meanings presents only weak evidence for a prospective *on* ASC in English. It is merely based on the idea that similar concrete meanings often develop into similar abstract, or more grammatical meanings, an idea that is part and parcel of grammaticalization studies (cf. e.g. Heine, Claudi & Hünnemeyer 1991).

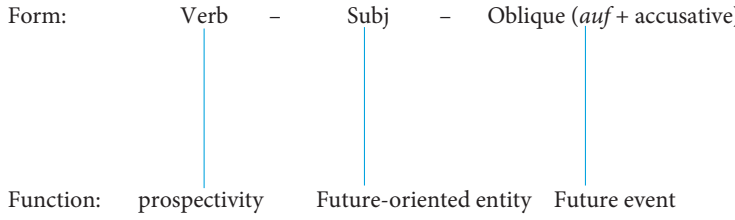

Figure 3. Prospective *auf* ASC

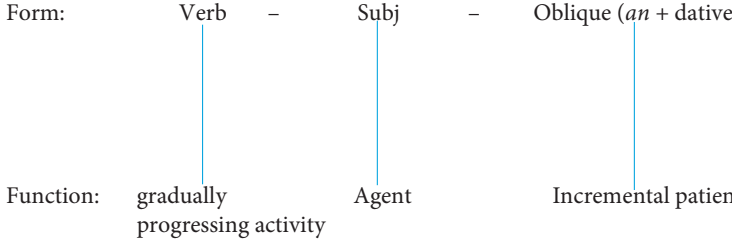

Figure 4. Incremental *an* ASC

See Rostila (in press: Section 2.3) for more details and discussion regarding the structure and semantics of both constructions.

4. Improved solution

The comparison with German has so far yielded three ASC candidates for English, each with the surface manifestation *on*, that could jointly generate the cases that Goldberg (2014) ascribes to her two senses of the Rely On construction. The purpose of this section is to show how and why the alternative solution might be more plausible. In the following, the properties of the proposed three constructions are recapitulated and their interrelationships are surveyed. Israel's (1996) findings regarding the development of polysemy in the history of the English *way* ASC are used as a basis for modeling these relations. This seems justified for at least two reasons: first, the emergence of polysemy is primarily a diachronic process, and hence it seems fruitful to exploit a diachronic parallel case to inform the synchronic description of the Rely On domain. Second, and more importantly, a comparison of the experimental results presented in Tomasello (2003) and Goldberg (2006) regarding the ontogeny of ASCs and Israel's (1996) empirical historical findings suggests that the emergence of ASCs both in ontogeny and phylogeny involves a generalization process catalyzed by high frequency items and

intermediate generalizations resulting from type frequency effects; see Rostila (in press: Section 3.3) for discussion. Therefore, the development[15] sketched by Israel (1996) for the *way* ASC might be at least roughly representative of the development of ASCs in general and hence serve as a legitimate model for the interrelationships of the Rely On constructions.

Here is an overview of the ASCs of the alternative solution and their interrelationships:

ASC 1: Signified by *on*; expresses incremental progress in ingestion; combines with relatively concrete verbs of ingestion like *nibble, chew*. In light of Israel (1996), it seems plausible that this ASC has gradually spread from such verbs to more abstract verbs of ingestion like *feed, prey* and *live*. Such verbs focus more on the aspect of sustenance in ingestion than progress in it,[16] and the occurrence of several such more abstract verbs in ASC 1 may have led to the emergence of an intermediate generalization, i.e. ASC 2, expressing sustenance – cf. Israel (1996: 223) for a similar process in the development of the *way* ASC, where "analogical extensions" to further verbs lead to "clusters of usage", which in turn license "more abstract schemas". The diachronic relationship between ASC 1 and ASC 2 amounts synchronically to a polysemy link between the two.

ASC 2: Likewise signified by *on*; expresses sustenance. This constructional semantics enables the use of verbs like *rely* and *depend* in the pattern, i.e. verbs expressing more abstract sustenance instead of 'sustenance by ingestion'. Some verbs of this type display clearly prospective/'hope' semantics, cf. *count, reckon, bet, call (on X to do Y)*.[17] By providing a semantically defined cluster and thus a basis for

---

15. Israel (1996) studies the development of the English *way* construction (e.g. *The wounded soldiers limped their way across the field*) on the basis of 1211 diachronic examples from the OED and 1047 contemporary examples from the OUP corpus, showing that the construction was extended to new verbs in mainly two ways: by analogy with individual verbs already occurring in it, and by generalizations across clusters of such verbs, the latter leading to semantically more radical extensions. Significantly, similar phenomena are identified by Tomasello (2003) and Goldberg (2006) in the development of ASCs in child language. This makes it a tempting hypothesis to assume that they characterize all development of ASCs – also that of the *on* constructions proposed in this chapter, which are partly motivated by this hypothesis.

16. See Detges & Waltereit (2002) for a study explaining grammaticalization with the aid of this type of switch between figure and ground. Since (contra Noël 2007) there are grounds for considering the emergence and generalization of ASCs a process of grammaticalization (cf. Rostila 2005, 2007, 2014, in press: Section 3), it is to be expected that such a potential general feature of grammaticalization processes should appear in the development of an ASC.

17. The FrameNet description of the Reliance frame (cf. <https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Reliance>; 23.06.2016) suggests that prospectivity is

a generalization (again cf. Israel 1996: 223), their occurrence in the construction may have led to the development of a further ASC, i.e. ASC 3, whose semantics prospectivity dominates.[18]

ASC 3: Signified by *on*; expresses prospectivity. Notably, gaining sustenance is not part of the semantics of this ASC. This has the advantage of enabling ASC 3 to generate a wider range of cases than Goldberg's extended sense, e.g. the following:

(4) a. Foreign investors are waiting on election results there.[19]
b. I insist on your being present.
c. She was intent on pursuing a career in business.
d. They decided on their course of action.

Cases like this are prospective in that the argument marked by *on*, the focus of the subject argument's attention, is to be understood as a future event; see Rostila (in press: Section 2.3) for details.

Notably, ASC 3 can also generate the examples that Goldberg (2014) gives of her extended sense, i.e. cases like *bet on Y, call on X to do Y*. In such cases, the construction contributes the element of prospectivity/'hope', while 'gaining sustenance', if present at all (cf. note 18), conceivably stems from verb semantics. The same goes for cases like *count/reckon on X*, which Goldberg (2014) does not exemplify.

Now, at least on a cursory look, it might seem that Goldberg's Rely On construction – i.e. a generalization over the two subsenses assumed by her – could generate all the relevant cases: when combined with verbs of concrete ingestion, it would provide the semantic component 'gaining sustenance' that is present in the background; with verbs of sustenance like *rely* and *depend* its meaning would overlap with verb semantics; and with prospective verbs like *bet* and *call* it would again contribute the element 'gaining sustenance'. However, the construction

---

always part of a reliance relation in that such a relation projects a Means action needed by the subject of the relation. Such an action can only follow the rise of the need for it; hence, a reliance relation seems inherently prospective. The presence of a prospective component in verbs like *rely* and *depend* in fact only lends support to the proposed analysis, since it can further explain why more overtly prospective verbs like *bet* appear in the same pattern: such an analogical extension may have been enabled by both a shared component of sustenance and that of prospectivity – or just by the latter, in case verbs such as *bet* do not exhibit a component of sustenance, as one referee of this chapter is inclined to assume (cf. also note 18).

18. In fact, some such verbs, e.g. *bet*, may have found their way into ASC 2 only on the basis of sharing the component 'prospectivity' with verbs such as *rely* and *depend* – i.e. they may not display the component 'sustenance' at all. Precisely this on the other hand may have led to their forming a cluster providing the basis for the development of ASC 3.

19. There probably also exists in present-day English a competing prospective ASC based on the preposition *for*, cf. *I'll wait for you/hope for the best/prepare for the worst* and Rostila (2015: 43).

would not be able to generate cases like (4): the element 'gaining sustenance' imposed by the construction is not part of their semantics. Since the three ASCs proposed here can generate all the cases considered, *and* those in (4), while at the same time being semantically more specific and hence psychologically more realistic, they seem to form the preferable option. To elaborate on the latter aspect: while Goldberg's (2014) two constructions seem to allow for a generalization that slumps them together, the three ASCs proposed here each display different semantics that cannot be generalized over. Thus, even apart from the number of ASCs employed, my solution is on a more specific level than Goldberg's. To the extent that knowledge of language is more likely to be item-specific than generalizable,[20] this seems to be a desirable feature of my solution. A further argument in favor of my solution is that the diachronic emergence and polysemy relations of the constructions proposed here seem viable in the light of the findings of Israel (1996), whereas the polysemy relation envisaged by Goldberg (2014) between the two subsenses of her Rely On construction is essentially *ad hoc*.

It would thus seem that Goldberg's solution can indeed be improved. The alternative solution captures more data – Goldberg's solution would actually seem to have an undergeneration problem, instead of simply overgenerating, as might be expected of a maximally general solution. Furthermore, my solution is motivated by principles probably applying to both the diachronic and the acquisitional development of ASCs. However, there is a caveat to be made here. First, the semantic considerations that my solution rests on are somewhat lax in that they are based on intuition, not on semantic test procedures. Second, undergeneration and overgeneration issues are less than clear in the domain of prepositional objects, where competition between different patterns with a low degree of generalization and preemption effects are rampant (cf. Rostila 2007: 197–204, in press: Section 2.2). The pros and cons of constructions proposed for this domain can therefore only be reliably assessed as part of a larger solution capturing polysemous and partially synonymous ASCs based on PO prepositions.[21]

---

20. If usage-based linguistics (see e.g. Bybee 2006) is right to claim that exemplars are stored along with generalizations that can be drawn from them, and Tomasello's (2003) view of language acquisition progressing from item-based categories to generalizations (while often remaining on an item-specific level) is on the right track, this would seem a legitimate conclusion: exemplars are primary for language, generalizations seem like an optional extra.

21. A great step in this direction is taken by Uhrig and Zeschel (2016), who propose several ASCs for the Rely On and neighboring semantic domains both in English and German on the basis of corpus evidence. Furthermore, they link these ASCs to frames and image schemas. However, they do not relate their proposal to that of Goldberg (2014) yet. Unfortunately, their paper came to my knowledge too late for me to be able to formulate here a synthesis of their insights with those of Goldberg (2014) and mine.

5. Concluding remarks

The analysis has shown that it may be useful to take a look at semantically similar constructions in another language when deciding how to split a certain domain of the target language into constructions. Such comparisons may function as eye-openers that show the possibility of alternative analyses. However, the viability of analyses inspired by such comparisons must be based on target language data and independent principles of what kinds of constructions are likely to exist. As regards the present case, the advocated solution is motivated by generating more relevant cases than Goldberg's approach, by being less general and thus psychologically more plausible, and by being based on empirical diachronic observations possibly representative of the emergence of ASCs in general. An obvious next step would be to compile a truly representative list of verbs occurring in the three constructions proposed and to analyze whether the proposed interplay between verb and construction semantics holds true for all of them. In addition, the history of the ASCs of the Rely On domain would be a fruitful subject of study, since a verification of the parallels proposed here to the history of the *way* ASC would go a long way towards showing that there are generalizable features in the diachrony of ASCs. Last but not the least, Rudanko's (1989)[22] careful analysis of the complements of verbs like *depend, rely* and *count* – conducted in true CxG spirit, since both form and function are considered – should be used as a basis for their CxG description, and the interplay of such constructions with the ASCs proposed here should be studied.

References

Boas, Hans C. 2003. *A Constructional Approach to Resultatives*. Stanford CA: CSLI.

Boas, Hans C. (ed.). 2010. *Contrastive Studies in Construction Grammar* [Constructional Approaches to Language 10]. Amsterdam: John Benjamins. doi:10.1075/cal.10

Boas, Hans C. 2011. Coercion and leaking argument structures in Construction Grammar. *Linguistics* 49(6): 1271–1303. doi:10.1515/ling.2011.036

Boas, Hans C. 2014. Lexical and phrasal approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40(1–2): 89–112.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4): 711–733. doi:10.1353/lan.2006.0186

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: OUP. doi:10.1093/acprof:oso/9780198299554.001.0001

22. Back then, I wasn't even a student of Juhani's yet; if he hadn't impressed me with his teaching in the 90's, I wouldn't be in linguistics today.

Detges, Ulrich & Waltereit, Richard. 2002. Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(2): 151–195. doi:10.1515/zfsw.2002.21.2.151

Goldberg, Adele E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5): 219–224. doi:10.1016/S1364-6613(03)00080-9

Goldberg, Adele E. 2006. *Constructions at Work. The Nature of Generalization in Language*. OUP: Oxford.

Goldberg, Adele E. 2014. Fitting a slim dime between the verb template and argument structure construction approaches. *Theoretical Linguistics* 40(1–2): 113–135.

Heine, Bernd, Claudi, Ulrike & Hünnemeyer, Friederike. 1991. *Grammaticalization: A Conceptual Framework*. Chicago IL: University of Chicago Press.

Hilpert, Martin. 2014. *Construction Grammar and its Application to English*. Edinburgh: EUP.

Israel, Michael. 1996. The *way* constructions grow. In *Conceptual Structure, Discourse and Language*, Adele E. Goldberg (ed.), 217–230. Stanford CA: CSLI.

Müller, Stefan & Wechsler, Stephen. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2): 1–76. doi:10.1515/tl-2014-0001

Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14: 177–202. doi:10.1075/fol.14.2.04noe

Perek, Florent. 2015. *Argument Structure in Usage-Based Construction Grammar* [Constructional Approaches to Language 17]. Amsterdam: John Benjamins.

Rostila, Jouni. 2005. Zur Grammatikalisierung bei Präpositionalobjekten. In *Grammatikalisierung im Deutschen*, Torsten Leuschner, Tanja Mortelmans & Sarah De Groodt (eds), 135–166. Berlin: Walter de Gruyter. doi:10.1515/9783110925364.135

Rostila, Jouni. 2006. Construction Grammar as a functionalist generative grammar. In *At the Crossroads of Linguistics Sciences*, Piotr P. Chruszczewski, Michał Garcarz & Tomasz P. Górski (eds), 365–376. Cracow: Tertium.

Rostila, Jouni. 2007. Konstruktionsansätze zur Argumentmarkierung im Deutschen. PhD dissertation, University of Tampere. <http://urn.fi/urn:isbn:978-951-44-7085-1>

Rostila, Jouni. 2014. Inventarisierung als Grammatikalisierung: Produktive Präpositionalobjekte und andere grammatikalisierte Linking-Muster. In *Grammatik als Netzwerk von Konstruktionen. Sprachwissen im Fokus in der Konstruktionsgrammatik*, Alexander Lasch & Alexander Ziem (eds), 127–153. Berlin: De Gruyter. doi:10.1515/9783110353693.97

Rostila, Jouni. 2015. Inside out: Productive German prepositional objects as an example of complements selecting heads. In *Perspectives on Complementation. Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 34–51. Houndsmills: Palgrave Macmillan.

Rostila, Jouni. 2016. Zur Integration von Argumentstrukturkonstruktionen in das *Historisch syntaktische Verbwörterbuch*. In *Historisch syntaktisches Verbwörterbuch. Valenz- und konstruktionsgrammatische Beiträge*, Jarmo Korhonen & Albrecht Greule (eds), 261–276. Frankfurt: Peter Lang.

Rostila, Jouni. In press. Argument structure constructions among German prepositional objects. In *Constructional Approaches to Syntactic Structures in German* [Trends in Linguistics. Studies and Monographs], Hans C. Boas & Alexander Ziem (eds). Berlin/New York: Mouton de Gruyter.

Rudanko, Juhani. 1989. *Complementation and Case Grammar. A Syntactic and Semantic Study of Selected Patterns of Complementation in Present-Day English*. Albany NY: State University of New York Press.

Schøsler, Lene. 2007. *The status of valency patterns*. In *Valency: Theoretical, Descriptive and Cognitive Issues*, Thomas Herbst & Katrin Götz-Votteler (eds.) 51–65. Berlin: Mouton de Gruyter.

Tomasello, Michael. 2003. *Constructing a Language*. A Usage-Based Theory of Language Acquisition. Cambridge MA: Harvard University Press.

Uhrig, Peter & Zeschel, Arne. 2016. *On means auf*: The semantics of English and German prepositional complements in contrast. Paper presented at the 9th International Conference on Construction Grammar, Juiz de Fora, 5 October.

Van Pottelberge, Jeroen. 2004. *Der am-Progressiv: Struktur und parallele Entwicklung in den kontinentalwestgermanischen Sprachen*. Tübingen: Narr.

Visser, Frederik. 1973. *An Historical Syntax of the English Language*, Part Three, Second Half: *Syntactical Units with Two and with More Verbs*. Leiden: E. J. Brill.

PART 2

Variation and change in complementation patterns

Aspects of the use of the transitive *into -ing* pattern in New Zealand English

Paul Rickman & Mark Kaunisto
University of Tampere

This chapter investigates the use of the transitive *into -ing* pattern in one of the world's youngest dialects of English, New Zealand English (NZE). We draw on evidence collected from a diachronic NZE corpus of newspaper English, which comprises 100 million words from the years 1995–98 and 2010–12. We categorise the matrix verbs according to semantic orientation, and provide comment on any change evident between the two time-frames of the corpus, and we identify several previously undocumented matrix verbs in this pattern. Comparison is drawn to American English in an attempt to determine similarities in the use of this surprisingly productive pattern.

Keywords: New Zealand English; complementation; transitive *into -ing* pattern; corpus linguistics

1. Introduction

The transitive *into -ing* pattern, exemplified by the sentence *Philip tricked Phyllis into paying for dinner*, has been given a thorough treatment in recent years (see, e.g. Davies 2012; Kim & Davies 2016; Rudanko 2000, 2002, 2005, 2006, 2011, 2015a, 2015b; Rudanko & Luodes 2005; Wulff et al. 2007).[1] This surge in attention parallels a rise in the general frequency of the pattern in the main varieties of English; something that has been convincingly demonstrated in the literature (see, e.g. Rudanko 2005, and Davies 2012 for figures). Despite all the recent coverage, the usage of the pattern in the young dialect of New Zealand English remains, to the best of our knowledge, unexplored. Should the best of our knowledge prove faulty, however, the diachronic approach afforded by a corpus created specifically for studies on NZE requiring large amounts of data, such as the present chapter,

1. As regards terminology, here we use 'transitive *into -ing*', as coined by Juhani Rudanko.

DOI 10.1075/slcs.195.05ric

will at least offer new insights into the use of the transitive *into -ing* pattern in this lesser studied variety of English.

With this study we aim to address the questions of how the frequency in NZE generally compares to other dialects; whether there is any apparent change in the qualitative or quantitative uses of the pattern in NZE over the time frame investigated, and whether NZE shows any innovative aspects of its use. With regard to the last point, the expectation that some kind of innovative usage will be found almost goes without saying, as each of the southern hemisphere Englishes is well known for lexical idiosyncrasies (Turner 1966: 112ff.; Bauer 1994; Hay et al. 2008: 65ff.), and NZE speakers are almost certain to add some local flavour to such a productive pattern.

This chapter is structured as follows: Section 2 introduces the pattern, covers the necessary theoretical aspects of transitive *into -ing*, and discusses the relevant points that have arisen from earlier research. Section 3 presents the corpus used in this study and methodological issues; Section 4 comprises the results and discussion, and Section 5 provides the summary and concluding remarks.

2. Background and earlier research

As earlier research has noted, a compelling property of the pattern is its apparent ability to provide fertile ground for new and innovative language use, and the array of matrix verbs documented with the transitive *into -ing* pattern is striking, to say the least. The following example, taken from the Corpus of New Zealand Newspaper English (CNZNE), illustrates the pattern.

(1) They force the roading authorities into providing expensive new motorways. (CNZNE *Dominion* 1995)

The pattern manifests certain syntactic and semantic properties that require discussion at this point. Syntactically, it should be noted that in sentence (1), the matrix verb *force* selects three arguments: the higher subject NP, the higher object NP, and the lower clause introduced by the preposition *into*. The lower clause contains a null subject NP, an understood argument, which, in an expanded syntactic representation, would be represented by the null element PRO. The postulation of a null subject, while not uncontroversial, is in accordance with much of the literature on argument structure today, and is found in recent mainstream, influential grammars of English, as this excerpt from Huddleston & Pullum (2002: 65) shows: "[m]ost non-finite clauses have no overt subject, but the interpretation of the clause requires that an understood subject be retrieved from the linguistic or non-linguistic context" (see also Brinton & Brinton 2010: 275). PRO obtains its

reference from the object of the higher clause, *the roading authorities,* and the sentence, therefore, is one of object control.

Verbs occurring in this pattern have been found to co-occur in some other patterns (Rudanko 2006; Kim & Davies 2016), the main two being the V + NP + *to*-infinitive pattern, which hosts three-argument verbs such as *force, persuade* and *pressure,* as in *They force the roading authorities to provide expensive new motorways*; and the V + NP pattern, which hosts two-argument verbs such as *fool, tease* and *tempt*, as in *They fool the roading authorities*. The third and less central pattern likewise hosts two-argument verbs, but in this case the second argument is not of the same semantic type as the previous pattern. The difference between the semantic roles of the NP object permitted with verbs like *fool/tease/tempt* versus those permitted with verbs like *talk*, is shown by comparing the awkward sounding (?)*they talked the roading authorities* with the more acceptable *they talked business*. Verbs such as *talk* form the third "very limited set of verbs" (Rudanko 2006: 315).

Semantically, it is important to note that the arguments of sentence (1) are assigned the semantic roles of Agent, Patient, and Goal, for the higher subject, higher object, and lower clause, respectively. Examples of superficially similar patterns are plentiful in the data, such as sentence (2); these do not fit the bill semantically or syntactically: the semantic role of the higher object has to be seen as Instrument, rather than Patient; and the understood subject of the lower clause obtains its reference from the subject of the higher clause, making this a case of subject, rather than object control. Tokens of this type fall outside the scope of the present study.

(2) She had put her life into building the bush reserve.
(CNZNE *Timaru Herald* 1997)

A connection between the transitive *into -ing* pattern and the Construction Grammar framework (Goldberg 1995) has been postulated in the literature; the pattern in this context is seen as a subtype of the caused motion construction. This helps to explain why so many unusual matrix verbs are so readily used in this pattern, and how speakers are able to understand the meaning without difficulty (for the original proposal of this, see Rudanko 2000: 81ff.; and for a recent detailed argument for the caused motion subtype, see Kim & Davies 2016). While we recognise the significance of the Construction Grammar approach, it will not be a focus of this analysis.

The history of the pattern has been traced with the help of large historical corpora. Using the Corpus of Historical American English (COHA), Davies (2012) provides a timeline for the development in AmE, and suggests that the pattern began life as an analogy of the *into* NP pattern, with the strongly nominal verb

being first entering into that pattern, as in *somehow I inadvertently willed it into being*, and providing a foothold for other verbs. Davies argues that "[r]ather than have the construction created "ex nihilo", it apparently started where [V NP into V-ing] would be least noticeable—where [into V-ing] could also be analysed as a noun, as with the pre-existing [into N]. And then once the [into V-ing] construction was firmly "established" in about the 1850s, the percentage of *being* decreased markedly" (Davies 2012: 166). Rudanko (2015a) traces the development in BrE using the Corpus of Late Modern English Texts (CLMET) version 3.0, and finds numerous examples of the fully fledged transitive *into -ing* appearing as early as the mid 18th century, along with a few of the *into being* examples, so it seems clear that the pattern was already firmly established in BrE prior to the 19th century. This supports earlier work by Rudanko on the history of the pattern in BrE (e.g. Rudanko 2000).

It has also been observed that verbs occurring in the transitive *into -ing* pattern express negative causation, and that the construction has a negative semantic prosody. Wierzbicka (1998: 125) stated that "the set of main verbs that can be used in this construction is quite limited and gives a clear clue to this construction's meaning. Thus, one can not only "talk" someone "into" doing something, but also "trick" them, "manoeuver" them, or "push" them. On the other hand, one cannot "encourage" or "induce" someone "into" doing anything. […] the *into* construction takes verbs that either imply, or are at least compatible with, the idea of manipulation (tricking, manoeuvring, and the like)" (see also Hunston & Francis 2000: 102 on this point). On the other hand, Rudanko (2006), citing the influence of the Great Complement Shift (Rohdenburg 2006) – a major feature of which sees *-ing* complements spreading at the expense of *to*-infinitive complements – points out a number of verbs of unflavoured, rather than negative, causation that are found in the pattern. The argument cites the verbs *impel, induce, influence, lead, motivate, prompt* and *stimulate*, all of which are also compatible with the *to*-infinitive complement (Kim & Davies 2016: 25f. also address this issue).

Verbs found to occur in the pattern have been categorised under various different semantic groups by different authors. One such grouping, comprising seven categories, is given below, with example verbs provided in italics.

Deceiving or tricking – *deceive, trick, bamboozle, con*
Forcing or pressuring – *force, pressure, arm-twist, coerce*
Bullying or frightening – *bully, frighten, harass, panic*
Verbally persuading – *cajole, inveigle, persuade, sweet-talk*
Luring or tempting – *lure, tempt, entice*
Coaxing or stimulating – *coax, stimulate, prod, stir*
Other – *annoy, bounce, hijack, rush* (adapted from Rudanko 2000: 83)

With regard to the semantics of the matrix verb, it has been suggested that there is variation to be found between BrE and AmE in the type of verb used most frequently with the transitive *into -ing*. Wulff et al. (2007) suggest that BrE speakers tend to favour verbs denoting physical force, such as *force, bully* and *push*, while AmE speakers lean towards more orally persuasive verbs, such as *talk, persuade* and *coerce*. This suggestion, based on the results of statistical analyses of newspaper data, has since been reinforced by Kim & Davies (2016), on the basis of more stratified corpora.

We close this section with a brief outline of the target dialect of this study.[2] As far as native speaker varieties of English are concerned, New Zealand English is one of the world's youngest, with the formation period 1840–1880 (Trudgill et al. 2000).[3] It was during this time that settlers, mainly European, the largest single group of whom embarked in various parts of England, arrived in New Zealand in significant numbers. The large influx was due mainly to settlement schemes, and was additionally helped along by the discovery of gold. Prior to this, New Zealand was populated by fewer than 100,000 Māori and little more than several hundred Europeans.[4] From 1840 then, new-dialect formation processes began to take place (Gordon et al. 2004; Trudgill 2004), and by the early 20th century a distinctive new dialect was emerging, and various, mainly accent-related, aspects of it were being commented on, often unfavourably, in newspapers and other forums (for a good summary of early attitudes to NZE, see Gordon and Abell 1990).

The phonological and phonetic features of NZE have received much attention since those early years, but it has only been in recent decades that NZE has been seriously examined with the aim of documenting any distinguishing grammatical features, and it is on this area that the present chapter seeks to shed further light. The nature of the pattern also inevitably takes us across the borders of the fields of

---

2. The present chapter represents part of a larger research project being carried out by the first author into various aspects of complementation in NZE. The aim is to document areas in which NZE differs from other varieties of English in this respect, and to discern to what extent, if at all, the phenomenon of colonial lag (Marckwardt 1958: 80; Görlach 1987; Hundt 2009) is relevant.

3. Compare Australian English (1800–1840), and South African English (1820–1860) (Trudgill et al. 2000: 302).

4. Precise figures are not available for the Māori population, but Belich (1996: 178) estimates the number to have been around 86,000 in 1769, the year when Europeans first made landfall, and around 70,000 in 1840. As for Europeans, the same source (*ibid.*: 198) states that they numbered around 300 in 1830, and 2,000 in 1840.

vocabulary and word-formation – areas in which NZE creativity, to repeat a point made in the introduction, is known to excel.

The following section continues by outlining the practical and methodological aspects of the study, and discussing the corpus that provided the data.

3. Data and methods

Our data come from the recently compiled Corpus of New Zealand Newspaper English (CNZNE), a 100-million-word, part of speech tagged corpus comprising two sub-sections, 1995–98 (42.6 million words), and 2010–12 (58.5 million words).[5] Material from a total of 13 metropolitan and provincial newspapers contributes to the total word count. The 1995–98 sub-section is structured so as to allow comparison to the British National Corpus (BNC) newspaper sub-section (9,412,174 words). A detailed account of the compilation process of the CNZNE can be found in Rickman (2017), but a few details here are in order.

The CNZNE was compiled using material from the archives of Fairfax Media, accessed via The Knowledge Basket news and information archive service.[6] Fairfax is one of the largest media conglomerates in Australasia, and, within New Zealand, it controls publications that span the length of the country. In compiling the corpus, care was taken to ensure that major metropolitan newspapers from as many of the main centres as possible were included (New Zealand's four main centres[7] are traditionally understood as being Auckland, the capital Wellington, Christchurch and Dunedin), as well as provincial papers from as many different regions of the country as possible.[8] The obvious reason for this is that while it is generally acknowledged that NZE is a great

5. The CNZNE was compiled by the first author of the present paper. At the time of writing, it is not freely available.

6. <www.knowledge-basket.co.nz>

7. Many thanks to an anonymous reviewer, who pointed out that the comparatively smaller cities of Hamilton and Tauranga are now seen by Radio New Zealand and Television New Zealand as main centres, as they have been included in weather forecasts of the main centres for at least a decade now. Thus the 'four main centres' notion may be a thing of the past. It is still in use, however, as can be seen on the pages of the Government-managed online encyclopedia of New Zealand, Te Ara <http://www.teara.govt.nz>, and we continue to use the term here.

8. Rickman (2017) also addresses the issue of multiple versions of a single article that can surface among the various publications under single ownership, like the publications under Fairfax, and the problems this causes in corpus compilation.

deal more homogeneous than BrE and AmE, there is still known to be some regional variation (Turner 1966: 163ff.; Bauer & Bauer 2000; Hay et al. 2008: 95ff.; Calude & James 2011).

As well as texts from all the usual newspaper sub-genres, material obtained from the Fairfax archive also includes service information. Macalister (2001: 37) discusses service information in newspapers, defining it as "lists – sports results, tv programmes, share prices, weather forecasts, and so on". He chose to exclude it from his corpora, and the same has been done with the CNZNE, on the grounds that it does not help all that much in giving a clear picture of the aspects of NZE under investigation in the research project that the corpus was compiled for.

Naturally, corpus size is an all-important factor with a study of this type, and the entire 100 million word corpus was used. By today's standards, however, even 100 million words, for some research purposes, lies towards the more modest end of the scale, so in order to ensure that we were getting all that the corpus had to offer, we opted to use the simple search string [into *ing], using the concordancing software AntConc 3.2.4w, and then manually sort the relevant tokens from the roughly 5,600 total hits. The main reason for this all-inclusive approach is that, as is well known, no corpus tagging software can provide 100 per cent accurate results, and it was noticed during the preliminary search phase of this project that a search string designed to find words tagged as verbs followed by 0–4 intervening words, followed by into v_*ing, for example, was not going to achieve perfect recall.

Furthermore, it was noted by Rudanko (2015b: 81ff.) as part of his search for what he terms "radically innovative" matrix verbs, that words not traditionally recognised as performing a verbal function (i.e. words not listed as verbs in the OED, according to Rudanko's criteria) may nonetheless function as matrix verbs in the transitive *into -ing* pattern. Therefore, a search string aimed at identifying verbs may well miss all such radically innovative examples. With our aim of identifying innovative NZE usage, it is of course vital to retrieve all such tokens.

The coordination of matrix verbs, as in *parents have to push and cajole their children into doing their homework*, is reasonably common in the data, and needs to be resolved in a systematic way. We chose not to record both verbs occurring in such structures, but rather to follow Rudanko (2005: 175) in recording only the rightmost verb, in this case *cajole*, due to its proximity to the complement. Only in the case of established verbal units, such as *name and shame*, which occurs once in our data, were the verbs not separated; these were recorded as one verb.

Although the connection between the preposition and the *-ing* clause is tight, it is nonetheless possible to find material inserted here, and, in addition to the primary search string given above, additional searches were run to find such tokens. Examples of the results are given below.

(3) a. …I find myself forced into not panning Jon Bon Jovi's latest effort. (CNZNE *Sunday News* 1997)

b. …the Palestinians will be pounded into basically giving up on their rights and leaving. (CNZNE *Dominion Post* 2011)

The most common word – and it is only ever a single word in our data – occurring in this position is *not*, which is found eight times in the entire corpus. Other words found in this slot include: *again, almost, basically, illegally, regularly*. In total 21 tokens of this type were retrieved.

One aim of this analysis is to identify new matrix verbs in the target pattern, and to this end we cross-checked our results against those of a selection of prominent sources on the same topic – Bridgeman et al. (1965), Francis et al. (1996), Kim & Davies (2016),[9] and Rudanko (2005). Any verbs not documented in any of these sources were further cross-checked against the relevant entries in the *Oxford English Dictionary* (OED), and the *New Zealand Oxford Dictionary* (NZOD). If the examples supplied in these two reference works did not include the verb used in the transitive *into -ing* pattern, the verb was then deemed eligible to be put forward as an example of innovative usage. The results of this process are presented in the following section, along with other results.

4. Results

This section presents the main observations on the uses of the transitive *into -ing* pattern in the two parts of the CNZNE. First we will pay attention to the overall frequencies in the use of the pattern and the most common matrix verbs used with the pattern, followed by a section observing verbs found in the data which have been previously unattested as selecting the complement pattern.

4.1 Type and token frequencies of the transitive *into -ing* pattern in the CNZNE

On the whole, the transitive *into -ing* pattern was quite frequent in the corpus data examined, with a total of 1,485 relevant tokens found. The token and type frequencies of the matrix verbs are given in Table 1.

9. We would like to thank Mark Davies for giving us access to the data that he and Jong-Bok Kim collected for their 2016 study, since only a small percentage of the large number of verbs they identified were listed in the actual paper.

Table 1. The frequencies of transitive verbs selecting *into -ing* complements in two sections of the Corpus of New Zealand Newspaper English (pmw = per million words)

CNZNE 1995–98 (42.6 million words)	**CNZNE 2010–12 (58.5 million words)**
635 tokens (14.9 pmw)	850 tokens (14.5 pmw)
134 different matrix verbs	145 different matrix verbs
190 different matrix verbs in total	
89 matrix verbs common to both sets	

As can be seen in Table 1, the normalized token frequencies of the pattern in the two periods were fairly similar, amounting to 14.9 instances per million words in 1995–98, and 14.5 per million words in the 2010–12 data. The type frequencies of the relevant transitive matrix verbs in the two periods were likewise close to one another, with 134 verbs found in the earlier period, and 145 verbs found in the latter. There was some variation as regards the sets of the matrix verbs with *into -ing* complements in that 45 verbs were unique to the 1995–98 period, 56 verbs were unique to the 2010–12 period, and 89 verbs were found in both sets.

Considering the normalized frequencies of the matrix verbs in the CNZNE data, some interesting observations can be made on them with regard to the findings in Davies (2012) and Rudanko (2015b). Firstly, in his study on the diachronic developments in the use of the pattern in COHA, Davies (2012) noted that in 2000–09 the normalized frequency of the pattern in the AmE data was 13.9 instances per million words, which is remarkably close to the frequency in the NZE data. Another point perhaps worth observing has to do with the slight decrease in the frequencies of the pattern in the two sets of NZE data. Although the drop in the normalized frequencies (14.9 to 14.5 instances per million words) may not be striking as such, in his study of the pattern in the TIME corpus, Rudanko (2015b: 82–83) also perceives a decline in the use of the pattern from the 1990s to the 2000s, with the 1970s being the peak period in its popularity.[10]

10. Interestingly enough, the normalized frequencies of the transitive *into -ing* pattern in the last couple of decades of the TIME Corpus are notably higher than those found in COHA and the CNZNE, as the frequencies in the TIME Corpus were 29.5 and 23.1 instances per million words in the 1990s and the 2000s, respectively (Rudanko 2015b: 82). It must be noted here that the three corpora are different as regards the text types or registers represented, and it is possible that the different frequencies observed in the TIME Corpus are reflective of the use of the pattern in magazine texts.

However, examination of data with a longer time span might be needed to conclude whether there is indeed a declining trend in the frequency of the transitive *into -ing* pattern in NZE.

The most commonly occurring matrix verbs taking an *into -ing* complement were largely the same in the two sets of data. In fact, the top 15 verbs were the same in both periods (namely *bully, coax, coerce, con, draw, dupe, fool, force, lure, mislead, pressure, provoke, push, talk,* and *trick*), with only slight differences in the order of frequency of the verbs in the two periods. Since the difference in time between the two periods studied is roughly that of half a generation, the similarity among the most prominent matrix verbs could be expected.

As regards the semantic characteristics of the verbs, the same major categories can be perceived as in the earlier studies of the pattern. There are instances where the verb involves different forms of pressure; for example, intimidation or surprise (e.g. *badger, blackmail, browbeat, jolt, kickstart, panic, shock*), or the use of physical force (e.g. *beat, bludgeon, bulldoze, stampede, thrust*). Persuasion (usually verbal) is another prominent semantic group (e.g. *cajole, convince, court, flatter, seduce, smooth talk, woo*), as well as deception (e.g. *bamboozle, confuse, deceive, hoodwink*). A less prominent semantic group in the data included words involving the idea of guidance or inspiration (e.g. *coach, encourage, enthuse, inspire, guide, lead, steer, support, usher*), which also serve as examples of verbs that carry a positive rather than negative sense.

4.2 Innovative usages

The productivity of the pattern is evident also on the basis of the findings in the CNZNE data, as altogether 19 innovative usages were found in the material, that is, 19 matrix verbs which have not been previously observed as having been used in the transitive *into -ing* pattern. In their study, Kim & Davies (2016) identified 335 innovative usages of the pattern in present-day AmE and BrE corpora (the Corpus of Contemporary American English [COCA], the BNC, and the US and UK sections of the Corpus of Global Web-based English [GloWbE]) containing in total as many as 1.32 billion words. From that perspective, finding innovative usages of the pattern in the NZE data is further testimony to how productive the pattern is today.

A particularly interesting case among the 19 innovative matrix verbs is the verb *monster*, the single instance of which is presented in example (4):

(4) … Mr Brownlee – who spent the week beavering away with officials while Mr Key was in Australia – had been monstered into bringing forward an announcement by the Beehive's ninth floor … (CNZNE *Dominion* 2011)

The use of *monster* as a verb is seen only in some varieties of English, and the OED and the NZOD list the verbal use with the regional label "Australian". This verb can thus be regarded as a genuinely new entry among the verbs selecting the *into -ing* pattern, arising from the AusE and NZE varieties. Another verb in this respect is *heavy*, which is listed in the OED as a verb, but not with argument patterns that would match that of the transitive *into -ing* pattern. The NZOD does include the word as a transitive verb, but no example is given of the verb with an *into -ing* complement. The example of the verb in the CNZNE is given in example (5).

(5) VEHICLE franchise war has been declared in Taranaki, with a major car manufacturer heavying its New Plymouth dealership into relinquishing an opposition franchise. (CNZNE *Daily News* 1997)

Of the 19 innovative matrix verbs found in the CNZNE data, three verbs comply with Rudanko's (2015b) definition of a "radically innovative" verb, i.e. they were not listed in the OED as verbs: *hard-talk, peer pressure,* and *turbo-shock*, illustrated in examples (6a)–(c):

(6) a. In a statement at James' hearing he said he was peer pressured into taking cocaine on a night out with friends and said it was a mistake. (CNZNE *Waikato Times* 2010)

b. The original version was popular Down Under but Sabbath apparently had to be hard-talked into playing it live because they didn't like having to use a piano and Mellotron (a type of keyboard) on stage. (CNZNE *Waikato Times* 2010)

c. … the experience of being hammered for pretty much eight months in a row seems to have turbo-shocked ministers into pulling up their socks in recent weeks. (CNZNE *Press* 2012)

The verbs *peer pressure, hard-talk,* and *turbo-shock* can also be regarded as innovative with respect to their structural property of consisting of two parts. In the case of *peer pressure* and *hard-talk*, the verbal use is probably a result of conversion or zero-derivation from a corresponding compound noun or an adjective + noun compound, while in *turbo-shock* the combining form *turbo* is attached to the base verb *shock*. In fact, morphological complexity is found in half of the innovations found in the corpus, as verbs involving compounding, conversion, prefixation, or the use as a phrasal verb are seen in seven other innovations of the pattern, namely *fast-track, force-feed, pressure-cook, stir up, re-energise, reinvigorate,* and *wrongfoot*. Here are illustrations of these verbs:

(7) a. Fast-tracking young graduates into teaching will result in disaster for many. (CNZNE *Dominion Post* 2011)
b. A European political elite had force-fed nations into accepting the euro, and despite multi-lateralism, it was still a world of nation states, he said. (CNZNE *Southland Times* 2012)
c. ARE horses getting their jumping tickets too easily these days? Are horses being pressure-cooked into jumping? A leading trainer posed these questions this week. (CNZNE *Evening Post* 1996)
d. Competition is a wonderful thing for stirring up local bodies into realising on assets. (CNZNE *Evening Standard* 1996)
e. … they are designed to revitalise and re-energise mature jobless people into wanting to get back into the work-force. (CNZNE *Press* 2011)
f. I want us to reinvigorate our families into taking up their full responsibility for their own. (CNZNE *Dominion Post* 2011)
g. So Bakugan are swamped by Gogos, who cede to Mighty Beans, whose time in the sun lasts just long enough to wrongfoot a grandparent into buying the present.[11] (CNZNE *Dominion Post* 2011)

The structural complexity seen in connection with innovative verbs was also noted by Rudanko (2015b: 85), and it may be that the innovativeness in word-forms correlates with the colourful and figurative usage characteristic of the language in magazines and newspapers.

The CNZNE data contained a further seven innovative matrix verbs with *into -ing* complements. These were the verbs *buoy, bustle, court, fuel, irk, link, mobilise,* with examples as follows:

(8) a. It was an effort which buoyed rider David Walsh into thinking he could win his second Wellington Cup … (CNZNE *Evening Post* 1996)
b. Canterbury looked a more focused unit early in the second half. Immediately from the kickoff, it bustled Southland into knocking the ball dead in goal and from the resulting drop out, hooker Malcolm Aldridge scored in the corner. (CNZNE *Southland Times* 1997)
c. But the immigrants themselves can hardly be blamed for that. They are courted into coming here as part of our immigration policy. (CNZNE *Sunday News* 1996)
d. … the story goes that the artist fell in love with a younger woman, but she did not fall in love with him. That fuelled him into painting his study of the female body with a grotesque beast lying on her stomach. (CNZNE *Nelson Mail* 2010)

11. *Bakugan, Gogos* and *Mighty Beans* are brands of children's toys, and the point here is that such toys enjoy only a brief period of popularity.

e. A cluster of real estate signs at the eastern entrance to Kurow has irked a resident into complaining to the community board representative. (CNZNE *Timaru Herald* 2011)

f. Girls' clothing is more aligned with fashion trends, linking girls early into being objects of the gaze and judgements of others. (CNZNE *Dominion Post* 2011)

g. … it was a matter of mobilising people into putting their concerns in writing. (CNZNE *Southland Times* 1997)

In some instances, the entries of the verbs in the OED and the NZOD may have included examples with *into* + NP or *to*-infinitive complements. The use of the verbs with these complements has become established before the *into -ing* pattern. Indeed, some of the instances of *into -ing* patterns in (8a)–(g) could be rephrased by using the other types of complements, some of them rather easily (e.g., *an effort which buoyed rider David Walsh to think that he could win his second Wellington cup; they are courted to come here; it was a matter of mobilising people to put their concerns in writing*). The V + NP + *to*-infinitive pattern, as noted in Section 1, is one of the main patterns with which many transitive *into -ing* verbs are also compatible.[12]

5. Summary and conclusion

This paper has investigated what could be seen as one of the most productive complement patterns in current English, with the aim of documenting its use and status in NZE. The data were drawn from a diachronic corpus of newspaper English with a time frame covering the better part of the last two decades. It was found that the overall frequency of the pattern does not change much over the half generation covered by the corpus, although the later section shows slightly lower frequency than the earlier. Further research would be needed to show whether this might be linked to Rudanko's (2015b) suggestion of a decrease in the use of the transitive *into -ing* pattern, as shown by data from the TIME corpus.

12. We agree with an anonymous reviewer, who pointed out that it would be highly relevant to compare the frequencies of the V + NP + *to*-infinitive pattern and the V + NP + *into -ing* pattern in NZE, among verbs that select both. While this falls outside the scope of the present study, we can refer to an observation by Juhani Rudanko on the verb *force* in AmE; it was noted that the results of two separate searches in COCA point to an "overwhelming predominance of the infinitival pattern in comparison to the *into -ing* pattern" (2015: 72). A complete appraisal of the situation in NZE will need to wait for future research.

The frequencies observed here are in line with those produced by Davies (2012), with data from COHA.

The data uncovered 19 matrix verbs that were previously undocumented in the transitive *into -ing* pattern, and at least two of these, *monster* and *heavy*, are attributed to local NZE lexical inventiveness. Clearly we are not claiming that the appearance of these verbs in this pattern in the CNZNE material represents the very first usage, only that these are possibly the first recorded examples in a study of this type. Among many of the innovative verbs, various word formation processes were seen to be actively used, and from this point of view, complex and creative matrix verbs such as *turbo-shock* will no doubt continue to appear, as the combination possibilities are almost endless.

The present analysis is based on newspaper language, and admittedly, the restriction to just one text type is not ideal. It should be remembered, though, that a large number of important studies have been conducted on the basis of newspaper material, and the genre does have its advantages, not least the wealth of different subject matters covered and the fact that it reflects current trends in language use. Provided the different types of newspapers – e.g. elite, popular, etc. – are represented in a corpus, valid inferences can be drawn from the results.

On a final note, it was something of a surprise that no influence of te reo Māori was evident among the matrix verbs in our material. We were pleased, however, to at least come across this example of the *into* NP pattern in the newspaper *The Press* from earlier years:

(9) Why, only ten years after a city's Jubilee, should the people be hakaed into another outburst, while a blank interval of forty years intervenes before the centenary celebrations? (*Press* 1928)

We have seen that the *into* NP pattern has a connection with the development of the transitive *into -ing* pattern, and, despite not having surfaced in our corpus data, it is quite possible that *haka* and/or other similar words from te reo Māori are being used in the *into -ing* pattern somewhere in New Zealand. Given the well-known negative attitudes towards te reo Māori during the early part of the 20th century, it seems that the author of this example was ahead of his/her time, but that *haka* did not succeed in making it into the ranks of verbs used with the *into -ing* pattern.

Acknowledgements

The authors would like to thank the audience of the 2015 Symposium on Complementation, held at the University of Tampere in honour of Professor Juhani Rudanko. Thanks must also

go to our two anonymous reviewers, whose suggestions have helped to improve the present work. We would also like to take this opportunity to acknowledge the scholarly achievements of Professor Rudanko, without which the present paper would not have been written.

References

Bauer, Laurie. 1994. English in New Zealand. In *The Cambridge History of the English Language*, Vol V: *English in Britain and Overseas, Origins and Developments*, Robert W. Burchfield (ed.), 382–429. Cambridge: CUP.

Bauer, Laurie & Bauer, Winifred. 2000. Nova Zelandia est omnis divisa in partes tres. *New Zealand English Journal* 14: 7–17.

Belich, James. 1996. *Making Peoples: A History of the New Zealanders from Polynesian Settlement to the End of the Nineteenth Century*. Harmondsworth: Allen Lane, The Penguin Press.

Bridgeman, Lorraine, Dillinger, Dale, Higgins, Constance, Seaman, David & Shank, Floyd. 1965. *More Classes of Verbs in English*. Bloomington IN: Indiana University Linguistics Club.

Brinton, Laurel J. & Brinton, Donna M. 2010. *The Linguistic Structure of Modern English*. Amsterdam: John Benjamins. doi:10.1075/z.156

Calude, Andreea & James, Paul. 2011. A diachronic corpus of New Zealand newspapers. *New Zealand English Journal* 25: 35–52.

Davies, Mark A. 2012. Some methodological issues related to corpus-based investigations of recent syntactic changes in English. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Traugott (eds), 157–174. Oxford: OUP.

Deverson, Tony & Kennedy, Graeme (eds.). 2005. *The New Zealand Oxford Dictionary*. Victoria: OUP.

Francis, Gill, Hunston, Susan & Manning, Elizabeth. 1996. *Collins Cobuild Grammar Patterns*, 1: *Verbs*. London: HarperCollins.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gordon, Elizabeth, Campbell, Lyle, Hay, Jennifer, MacLagan, Margaret, Sudbury, Andrea & Trudgill, Peter. 2004. *New Zealand English: Its Origins and Evolution*. Cambridge: CUP. doi:10.1017/CBO9780511486678

Gordon, Elizabeth & Abell, Marcia. 1990. This objectionable colonial dialect: historical and contemporary attitudes to New Zealand speech. In *New Zealand Ways of Speaking English*, Allan Bell & Janet Holmes (eds), 21–48. Clevedon: Multilingual Matters.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8: 41–60. doi:10.1075/eww.8.1.05gor

Hay, Jennifer, Maclagan, Margaret & Gordon, Elizabeth. 2008. *New Zealand English*. Cambridge: CUP. doi:10.3366/edinburgh/9780748625291.001.0001

Huddleston, Rodney & Pullum, Geoffrey. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? In *One Language Two Grammars?* Günter Rohdenburg & Julia Schlüter (eds), 13–37. Cambridge: CUP. doi:10.1017/CBO9780511551970.002

Hunston, Susan & Francis, Gill. 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English* [Studies in Corpus Linguistics 4]. Amsterdam: John Benjamins. doi: 10.1075/scl.4

Kim, Jong Bok & Davies, Mark A. 2016. The into-causative construction in English: A construction-based perspective. *English Language and Linguistics* 20: 55–83.

Macalister, John. 2001. Introducing a New Zealand newspaper corpus. *New Zealand English Journal* 15: 35–41.

Marckwardt, Albert. 1958. *American English*. Oxford: OUP.

Oxford English Dictionary, online. <http://www.oed.com> (February 2016).

Rickman, Paul. 2017. Aspects of verb complementation in New Zealand Newspaper English. In *Changing English: Global and Local Perspectives*, Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds), 169–190. Berlin: Mouton de Gruyter.

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166 Bern: Peter Lang.

Rudanko, Juhani. 2000. *Corpora and Complementation: Tracing Sentential Complementation Patterns of Nouns, Adjectives and Verbs over the Last Three Centuries*. Lanham MD: University Press of America.

Rudanko, Juhani. 2002. *Complements and Constructions*. Lanham MD: University Press of America.

Rudanko, Juhani. 2005. Lexico-grammatical innovation in current British and American English: A case study on the transitive into -ing pattern with evidence from the bank of English corpus. *Studia Neophilologica* 77: 171–187. doi: 10.1080/00393270500384791

Rudanko, Juhani. 2006. Emergent alternation in complement selection: the spread of the transitive 'into –ing' construction in British and American English. *Journal of English Linguistics* 34: 312–331. doi: 10.1177/0075424206293600

Rudanko, Juhani. 2011. *Changes in Complementation in British and American English: Corpus-based Studies on Non-finite Complements in Recent English*. Houndmills: Palgrave Macmillan. doi: 10.1057/9780230305199

Rudanko, Juhani. 2015a. Wheedled me into lending him my best hunter: Comparing the emergence of the transitive *into -ing* pattern in British and American English. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 128–140. Houndmills: Palgrave Macmillan.

Rudanko, Juhani. 2015b. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Houndmills: Palgrave Macmillan.

Rudanko, Juhani & Luodes, Lea. 2005. *Complementation in British and American English*. Lanham MD: University Press of America.

The Corpus of New Zealand Newspaper English (CNZNE).

Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Oxford: OUP.

Trudgill, Peter, Gordon, Elizabeth, Lewis, Gillian & Maclagan, Margaret. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36: 299–318. doi: 10.1017/S0022226700008161

Turner, George William. 1966. *The English Language in Australia and New Zealand*. London: Longman.

Wierzbicka, Anna. 1998. The semantics of English causative constructions in a universal-typological perspective. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. I, Michael Tomasello (ed.), 113–153. Mahwah NJ: Lawrence Erlbaum Associates.

Wulff, Stefanie, Stefanowitsch, Anatol & Gries, Stefan T. 2007. Brutal Brits and persuasive Americans. In *Aspects of Meaning Construction*, Günter Radden, Klaus-Michael Köpcke, Thomas Berg & Peter Siemund (eds), 265–282. Amsterdam: John Benjamins.
doi: 10.1075/z.136.17wul

Complementation of *ashamed* – diachrony and determinants of variation

Mikko Höglund
Stockholm University

The present chapter will contribute to the area of English complementation studies by investigating the adjective *ashamed*. The study focuses on the diachronic development of the use of *ashamed* during the past two centuries in different environments, and the variation between two rivalling non-finite complements, *to* infinitive and *of -ing*. The data was collected from COHA and COCA. The results show that since the 1870s *ashamed* has been used less and less, and that it occurs nowadays regularly without a complement whereas in the 19th century the *to* infinitive complement was common. The *of -ing* complement is not used much, but some contexts such as negation or copula in the complement clause were observed to attract this complement type.

Keywords: complementation; adjective; ashamed; corpus; diachrony

1. Introduction

Grammatical variation is an area of linguistics that has been much studied, and advancements in computer technology have brought about increasingly more efficient methods of corpus linguistics with which variation can be studied in more meticulous detail than before. As larger and larger corpora become available, it is possible to investigate more subtle factors behind variation as well as variation in infrequent items. One particular area that has benefited from the rise of corpus linguistics is complementation studies. By using large corpora it is possible to extract thousands of examples of complementation patterns, and uncover and quantify different variables that might contribute to complement variation. The majority of complementation studies in the past have focused on the complementation of verbs, and adjective complementation has received less attention. There are, however, studies that discuss adjective complementation, at least in the English language: e.g. Kjellmer (1980), Rudanko (1999, 2000, 2006, 2010, 2011, 2012, 2014, 2015), Van linden and Davidse (2009), Van linden (2010, 2012), Mindt (2011), Höglund (2014) and Havu & Höglund (2015).

DOI 10.1075/slcs.195.06hog

The present chapter aims to contribute to the field of adjective complementation studies by examining the adjective *ashamed* and the variation in its complementation. The study focuses on two patterns, the *to* infinitive complement (*ashamed to admit*) and the *of -ing* complement (*ashamed of admitting*).[1,2] The goal of the present paper is to provide evidence for two questions: which features affect the complement selection and have there been diachronic changes in the choice of the complement.

First, it has been shown that different features – syntactic, semantic or cognitive – can affect the choice of the complement when there is more than one option (e.g. complexity (Rohdenburg 1996), extraction (Vosberg 2003), agentivity (Rudanko 2010, 2017), animacy (Höglund 2014)). The main objective of the present paper is to analyze the environments in which the two complementation patterns occur, and analyze the factors that may affect the choice of one complement type over the other.

The second objective is related to the diachrony of the English complementation system. It has been observed that, in the realm of non-finite complementation, complements which include gerund forms have become more frequent at the expense of *to* infinitive complements in Late Modern English (e.g. Fanego 1997; Rudanko 2006, 2010, 2011, 2012; Rohdenburg 2006; Mair 2006 § 4.8.2; Vosberg 2003; Egan 2008; De Smet 2013). This phenomenon is called the Great Complement Shift (Rohdenburg 2006, Vosberg 2006), and it is hypothesized here that *ashamed* follows the same pattern – increasingly more *of -ing* complements at the expense of *to* infinitives. The aim is to chart the development of the complementation patterns from the 19th century until the present day and record the changes in their frequency of use. In order to investigate the complementation patterns of *ashamed*, material was collected from the Corpus of Contemporary American English (COCA, Davies 2008-) and the Corpus of Historical American English (COHA, Davies 2010-).

The present chapter is constructed as follows: Section 2 sets the stage for the present investigation by discussing complement variation, the meaning of *ashamed*, and reviewing previous studies on similar topics. Section 3 presents the materials and method, and Section 4 reports the findings of the corpus study together with the analysis of the results. Section 5 concludes the chapter.

---

1. According to Herbst et al. (2004), *ashamed* also occurs in the pattern *about -ing*, but it was excluded from the present study because it seems to be very rare (only 10 instances in COCA).

2. Rudanko (2017) discusses the same topic, and there is some overlap between Rudanko's work and the present chapter. It must be noted, however, that the two papers were authored independently and unknowingly of each other, and while Rudanko's paper focuses on the Choice Principle, the present chapter has a different focus and scope.

2. Background

It is a well-known fact that many predicates in English can occur with more than one type of complement. For instance, the verb *claim* may have a finite (1a) or non-finite (1b) complement:

(1) a. She claimed *that she was a linguist*.
b. She claimed *to be a linguist*.

In both (1a) and (1b) the basic event in the world that is described is the same, but the conceptualizations evoked by the different complement types are different.[3] In the same vein, some predicates, such as *start*, may have two non-finite complement options:

(2) a. She started *to conjugate verbs*.
b. She started *conjugating verbs*.

In (2a) there is a *to* infinitive complement, and in (2b) an *-ing* complement. Of the four examples in (1)–(2), it is clear that only in (1a), which has a finite clause complement, there is an overt subject in the complement clause. In (1b) and (2ab), which have non-finite complement clauses, there is no explicit subject in the complement clause, and it is understood, since there is only one participant in the sentence, that the subject in the main clause is also the primary participant in the complement clause event. That is, in (1b) it is the same person claiming and possibly being a linguist, and in (2ab) it is the same person that starts and conjugates.

Example (3) illustrates yet another case of complement variation, the main topic of the present chapter:

(3) a. He was ashamed *to answer* the question.
b. He was ashamed *of answering* the question.

In principle, these two complementation patterns are very similar to those in (2). In both pairs the main predicate retains more or less the same meaning regardless of the complement type, and in both the subject of the main clause is the primary participant also for the predicate in the complement clause. However, there is a structural difference between the (b) examples of (2) and (3). Whereas in (2b) the complement is a verb phrase (*-ing* clause), in (3b) it is a prepositional phrase

3. It is claimed that the non-finite pattern is semantically more integrated to the matrix clause than the finite pattern (Givón 1985, 1991, 2001; Rohdenburg 1995), and that the finite pattern is cognitively more explicit and easier to understand (Rohdenburg 1996 and references therein). This also harkens back to Bolinger's (1968: 127) generalization that "difference in syntactic form always spells a difference in meaning".

headed by the preposition *of*, and the *-ing* clause is a complement of the preposition. Despite the difference, the variations shown in (2) and (3) are considered to be of similar types since they both have the *-ing* form, and it can be assumed that they are subject to similar tendencies regarding diachronic change and factors that affect the choice of one or the other complement type.

The OED mentions both complement types, the *to* infinitive and *of -ing*, in the entry for *ashamed*. They both convey the meaning 'affected with shame' and the cause of shame is expressed in the complement (cf. senses 1 and 2). Moreover, in sense 3 the infinitive is said to have additional meanings: in sense 3a the idea of reluctance is introduced ('reluctant through fear of shame *to*') and sense 3b states '[w]ith negative: prevented or deterred by fear of shame from'. The following examples are given of senses 3a and 3b, respectively:

(4) He would have made us ashamed to show our Heads.
(Joseph Addison, *The Spectator*, 1711)

(5) He was not ashamed to answer that he could not live out of the royal smile.
(Thomas Babington Macaulay, *The history of England from the accession of James II*, 1849)

Considering Bolinger's (1968: 127) famous generalization "difference in syntactic form always spells a difference in meaning", it seems natural that one complement type has slightly different semantic scope than the other. However, the issue with sense 3a is that the assumed reluctance sense might not be there only with and because of the *to* infinitive. The reluctance meaning can be seen in (4), but it might be a general property of adjectives that are negative or have negative connotations. If someone has a negative feeling towards something, they are usually also reluctant in relation to acting on the matter. A sentence similar to (4) but with the *of -ing* complement, *He would have made us ashamed of showing our heads*, could convey the same reluctance as the original sentence with the infinitive complement in (4).[4] In the same vein, sense 3b, exemplified in (5), could be conveyed by an *of -ing* complement as well: *He was not ashamed of answering* […].[5] It seems to be a possibility that even if the *to* infinitive complement has the

4. One reviewer disagreed with this interpretation, saying that the *of -ing* version rather conveys the 'affected with shame' sense. The present author acknowledges that semantic assessments like this are very difficult to make, and also difficult to prove one way or the other. Be that as it may, at least for some speakers the reluctance is also present in the *of -ing* pattern.

5. In this particular example the following *that*-clause also has an effect on the complement choice, as "*ashamed to* V *that*" seems to be a somewhat fixed pattern. See discussion in Van linden (2012: 260–261), Rudanko (2017: 52–53) and Section 4.2 in the present chapter.

two additional meanings introduced in the OED sense 3, the *of -ing* complement is able to convey the same meanings. However, the meanings are very close to each other, and subtle nuances are quite difficult to tease apart, even with the help of corpus data.

Rudanko (2017) investigated the two complementation patterns of *ashamed* in detail, focusing on the so-called Choice Principle, which is introduced in his earlier work (2010, 2011), as a possible factor in the complement choice. The Choice Principle is formulated in the following way:

> In the case of infinitival and gerundial complement options at a time of considerable variation between the two patterns, the infinitive tends to be associated with [+Choice] contexts and the gerund with [–Choice] contexts.
>
> (Rudanko 2017: 20)

Broadly speaking, in [+Choice] contexts the subject participant is agentive, i.e. acts volitionally and has responsibility and control in the event/action denoted by the lower clause predicate, and in [–Choice] contexts the subject is non-agentive and has the semantic role of patient or theme, not being in control of the event. Rudanko has observed in the studies mentioned above that [–Choice] contexts more readily select the gerundial complement option (no matter whether it occurs by itself (example 2b) or as a complement to a preposition (example 3b)) than [+Choice] contexts. This observation is also put to test in the present chapter.

Regarding other factors that may have an effect on the choice of complement, Vosberg (2003) noted that when an element is extracted from its canonical position in the complement clause (e.g. by relativization or topicalization), the infinitival complement option tends to be favored. In addition, other factors that increase the complexity of constructions (passive, negation) have been observed to have an effect on the choice of complement (e.g. Rohdenburg 1996, 2015), and these will also be investigated in this chapter as possible factors affecting the selection of different complementation patterns with *ashamed*.

3. Materials and method

The data for the present study was collected from the Corpus of Contemporary American English (COCA, Davies 2008-) and the Corpus of Historical American English (COHA; Davies 2010-), which comprise approximately 520 and 400 million words, respectively. These corpora were selected because *ashamed* is not a very frequent item and large corpora were needed in order to get enough data about its use.

COHA was investigated in order to get an overview of the development of different complementation patterns with which *ashamed* occurs. Random samples of 200 tokens of *ashamed* were investigated from selected six decades (1830s, 1850s, 1900s, 1920s, 1980s and 2000s) and the different constructions recorded. This procedure gave information on the development and distribution of different complementation patterns throughout the past two centuries.

COCA was investigated in order to get synchronic material in order to investigate the present situation regarding the complementation of *ashamed* and the different features that may affect complement selection. The search string for the *to* infinitive complementation pattern was simply "ashamed to" and in order to find instances in which there was an intervening item, also "ashamed * to" was used. Precision would naturally have been better if the tag for the infinitival verb form had been added to the search strings, but in order to maximize recall the simple strings were used. A manual screening was done to discard irrelevant instances which the simple search strings found, such as (6a) in which *to* is followed by an NP. As can be seen in (6b), cases in which the *of -ing* complement would not be possible were also discarded, since there is no possibility for variation.

(6) a. I'm ashamed to this day that I have to talk about this. (COCA, spoken, 1996)

b. He's ashamed to be singing such childishness. (COCA, fiction, 1998)

In theory, the *of -ing* complement would be possible in (6b), *He's ashamed of being singing such childishness*, but in practice this is not used (no instances in COCA, COHA, GloWbE or NOW) and there is no variation. All in all, a small number of instances were disregarded: 24 out of 816, leaving the total of relevant instances at 792.

In order to find the *of -ing* pattern, the string "ashamed of *ing" was used. Here, as with the *to* infinitive pattern, possible intervening items were also taken into consideration by using two additional searches: "ashamed * of *ing" and "ashamed of * *ing". The tag for *-ing* forms was not used, instead *ing was used to maximize the recall. Again, a manual screening was done and 47 irrelevant instances were discarded, leaving the total number of relevant instances at 142.

After the relevant instances were identified, they were analyzed in terms of different features, and the aim was to find correlations between the choice of complementation pattern and the investigated features. As mentioned above, most of the features were selected on the basis of previous studies on complementation (e.g. Rohdenburg 1996, 2015; Rudanko 1999, 2010, 2014, 2015, 2017; Vosberg 2003) and what they found to be factors affecting the choice of complement (extraction, negation, agentivity), and some were included on exploratory basis in order to see whether they yield any relevant results (perfect aspect or copular construction in

the complement). Examples from COCA of the features with both complementation patterns are found in Table 1.

Table 1. Features investigated in sentences with *ashamed* (examples from COCA)

	***to* infinitive**	***of -ing***
Perfect aspect	*She's ashamed to have forgotten him.*	*Hannah was ashamed of having harbored such an unworthy thought.*
Copula	*I am ashamed to be an American.*	*For the first time in my life I am ashamed of being an American.*
Passive	*It broke my heart that he was ashamed to be seen with me.*	*Do you think Julian's ashamed of being raised without money, …*
Non-agentive	*I'm ashamed to have Bob Dole represent us because…*	*Shirley was ashamed of falling asleep and…*
Negation in complement	*I'd be ashamed not to have great bunches of pansies…*	*She was ashamed of not being able to enjoy it.*
Negation in main clause	*And I'm not ashamed to fight for my country!*	*One should not be ashamed of eating a substantial meal…*
Extraction	*I was able to ask those questions that I felt so ashamed to ask before.*	*…and the fact that he was even with Lucy was a rule he claimed he was ashamed of breaking.*

Each relevant instance in the data was checked for the features, and the results are presented and discussed in Section 4. As can be seen in Table 1, all the features are indeed found with both complementation patterns in the data set. This highlights the fact that the differences and tendencies, if there are any, are more quantitative than qualitative in nature.

4. Findings and discussion

The present section reports the results of the corpus investigation, and presents a discussion of the findings. In 4.1, a general overview of the development of *ashamed* in COHA is given, together with the distribution of all the different complement types with *ashamed*. This is followed in 4.2 by the results of the analysis of different features that could have an effect on the choice of the complement.

4.1 Development of *ashamed* and its complements in COHA

Before *ashamed* and the two complementation patterns that are in focus in this chapter are discussed, it is useful to have a look at the development of the use of the

lexeme *ashamed* and the distribution of all the different environments in which it occurs in order to provide a frame of reference for the results of the variation between *to* infinitives and *of -ing* complements. Figure 1 displays the normalized frequency numbers of *ashamed* during the past two centuries in COHA.

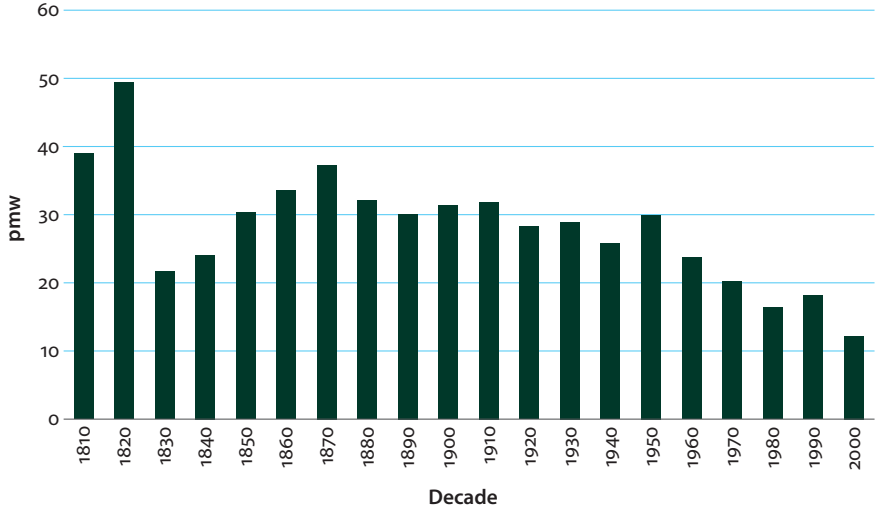

Figure 1. The use of *ashamed* in COHA

As can be seen in Figure 1, there has been some changes in the use of *ashamed* during the past 200 years. At first, it seems to have been very frequent in the 1810s and 1820s, but closer inspection reveals that in the 1810s data, 17 out of 42 instances come from one source, and in the 1820s data, 186 out of 342 hits are from only four writers. If those overrepresentations are discarded, the frequencies per million words drop to 21.17 and 22.52, respectively. Taken this into account, it looks like there is a steady increase in the use of *ashamed* in the first half of the 19th century, reaching a peak in the 1870s, and after that, there has been more or less a steady decrease decade by decade until the 21st century. Very similar declining trends towards the present day are seen for the two complementation patterns in Figures 2 and 3.

In Figure 2 the high frequencies per million words in the first two decades are again explained by idiosyncratic use: in the 1810s two writers are responsible for half of the instances, and in the 1820s three writers cover 42% of the hits. Both complementation patterns reach their peak in the 1870s when *ashamed* is used most frequently, and after that they show a declining trend, as is expected given the decreasing numbers of the adjective *ashamed* in general.

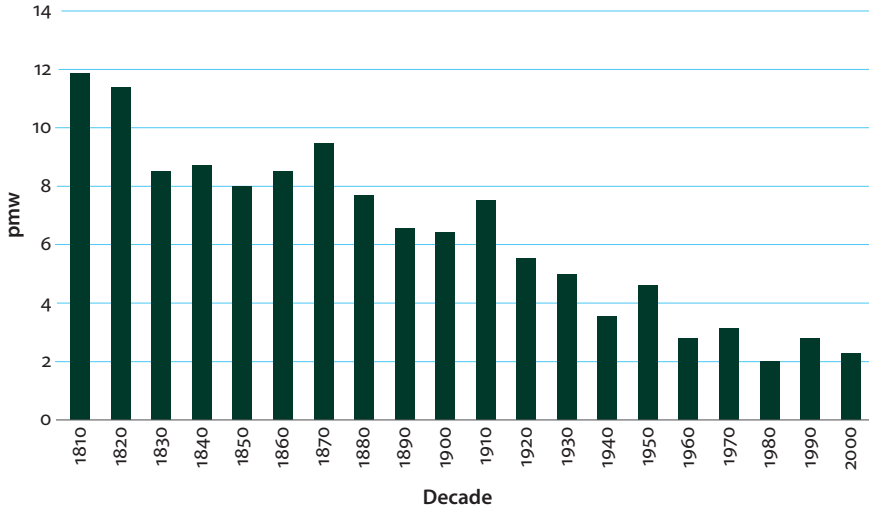

Figure 2. *Ashamed to* + infinitive in COHA (ashamed to _v?i*)

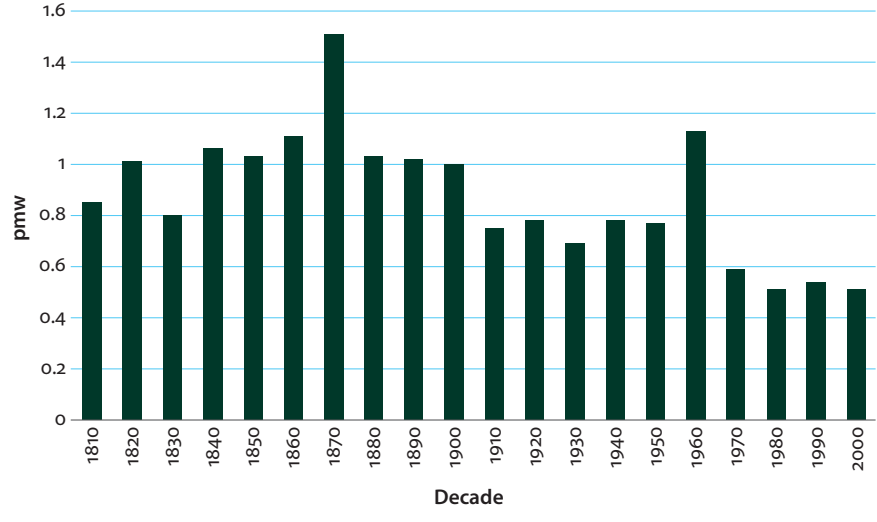

Figure 3. *Ashamed of -ing* in COHA (ashamed of _v?g*)

In order to shed more light on the development of the complementation patterns, the numbers need to be looked at in relation to the decreasing use of the adjective *ashamed* in general and also in relation to other complement types. In addition to the two complement types in focus in the present chapter, *ashamed* can occur independently without a complement, with a *that*-clause and with an *of* NP/clause complement. These are illustrated in (7):

(7) a. She felt ashamed and rebuked. (COHA, fiction, 1852)
b. We are ashamed that a taste so corrupt should so far prevail… (COHA, magazine, 1850)
c. Carey was never ashamed of the humbleness of his origin. (COHA, non-fiction, 1859)
d …to appear ashamed of what I glory in… (COHA, fiction, 1856)

As the purpose of this part of the investigation is to show the general development of *ashamed* from the 19th century to present day, it was considered sufficient to analyze samples from two decades in the 19th century, 1830s and 1850s, two decades from the beginning of the 20th century, 1900s and 1920s, and two recent decades, 1980s and 2000s. For each of the six decades a random sample of 200 instances of *ashamed* was investigated in terms of the complement type selected. The distribution of different complementation patterns with *ashamed* is illustrated in Figure 4, which shows the proportions (of the 200 tokens of *ashamed* per decade that were examined) of each pattern in which *ashamed* occurs.

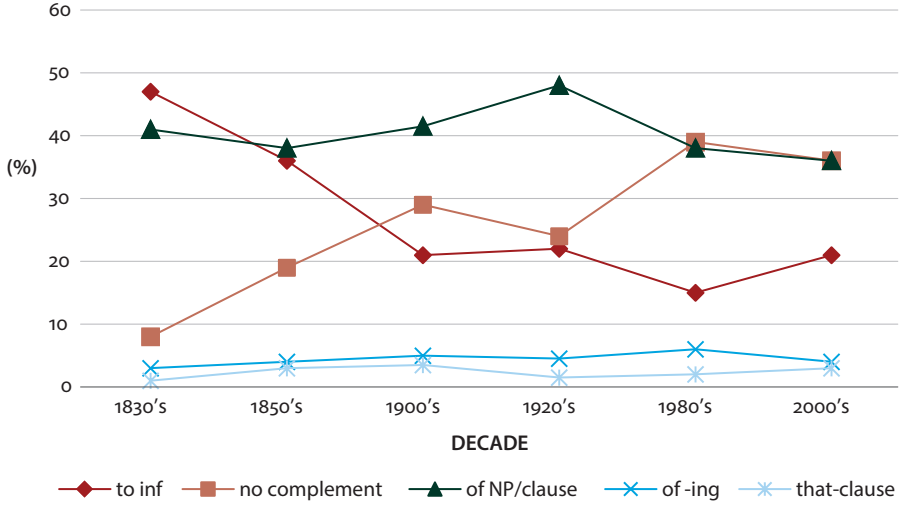

Figure 4. Development of the proportions of the complement types with *ashamed* in COHA

It looks like in the proportions of three of the patterns – *of* NP/clause, *of -ing* and *that*-clause – there have not happened any major changes during the past two centuries. The most interesting and greatest changes have happened in the proportion of *to* infinitives and the 'no complement' category. Since the 19th century, *ashamed* has been less and less frequently complemented by the *to* infinitive. Based on previous complementation studies dealing with the Great Complement Shift this can be expected, but there the reason for the decreasing use of the *to* infinitive is

observed to be the increase of the gerundial complement option. However, in this case the *of -ing* pattern has remained fairly marginal throughout the two centuries, and it seems that with *ashamed*, it is the 'no complement' category that has become more frequent. In these cases with no overt complement, the reason or cause for being ashamed is to be inferred from the context. It is difficult to say why in the 19th century it was more typical to overtly express the cause for being ashamed in the complement than today – perhaps the reason lies in stylistic changes in writing that have taken place since the 1800s.

To sum up, the use of the adjective *ashamed* has been in decline since the 19th century, and the use of the two complementation patterns under scrutiny, *to* infinitive and *of -ing*, have naturally undergone the same development. However, there have been some quite notable changes in the proportions of different complement types that occur with *ashamed*. Whereas in the 19th century the *to* infinitive and *of* + NP were by far the most frequent patterns with *ashamed*, today the *of* + NP pattern is accompanied by the complementless variant as the two most frequent patterns. The proportion of the *to* infinitive of all the instances of *ashamed* has dropped considerably since the 1830s to present day (from 47% to 21%), whereas the proportion of the complementless tokens has more than quadrupled (from 8% to 36%).

4.2 Choice between *to* infinitive and *of -ing* complement

This subsection presents the results from the analysis of different features introduced in Section 3 that possibly have an effect on the choice of complement. Altogether 934 relevant instances of *to* infinitives and *of -ing* complements were extracted from COCA and analyzed, of which *to* infinitives cover the larger proportion, 84.8%, with 792 tokens. This leaves *of -ing* at 15.2% with 142 tokens. As the *to* infinitive is by far the more frequent option, it begs the question which environments trigger the *of -ing* complement.

Given that all the features investigated occur with both complementation patterns, the possible differences are measured in quantitative terms. This can be done in two ways: we can look at absolute numbers or numbers relative to the overall frequencies of the two complement types. The former method is fairly simple: one feature at a time is considered and it is examined which complement type is favored in absolute numbers when that particular feature is present. Here it is expected that because of the much higher overall frequency of the *to* infinitive, it would also be the more frequent pattern with each feature. On the other hand, if the *of -ing* pattern is more frequent in absolute numbers with a certain feature than the *to* infinitive, then it is a very strong indicator that that feature strongly affects the selection of the *of -ing* pattern.

The latter way of measuring is to relate the proportions of the two complement types with each feature to the proportions of the two complement types overall (84.8% *to* infinitive and 15.2% *of -ing*). Here, if with a certain feature the proportions differ considerably from the overall proportions, it can be seen as an indicator that that particular feature may have an effect on the choice of the complement, of course depending on how much the proportions differ from the overall proportions.

Table 2 shows the numbers from the investigation in COCA and its 934 tokens of *ashamed* with the two complement types in terms of the features investigated. It must be noted that naturally there were tokens that did not have any of the features investigated, so the numbers do not add up to 934.

Table 2. Frequencies of the two complement types with different features in COCA

	***to* inf**	***of -ing***	**% of *of -ing* (overall 15.2)**
Perfect	17	9	34.6%
Copula	76	51	40.2%
Passive	34	8	19.0%
Non-agentive	56	13	18.8%
Negation in complement	2	7	77.8%
Negation in main clause	170	29	14.6%
Extraction	50	4	7.4%

As can be seen in Table 2, and as suggested by the overall frequencies of the two patterns, the *to* infinitive is the more frequent complement type with most features. There is only one exception: negation in the complement clause. The only two instances found in the data with the *to* infinitive, and two examples of the *of -ing* pattern are provided in (8) and (9), respectively:

(8) a. … something Richard was ashamed not to expect from the depths of a mobile home … (COCA, magazine, 1994)
b. … if I sold cut flowers in a shop I'd be ashamed not to have great bunches of pansies all spring. (COCA, magazine, 1991)

(9) a. Why be ashamed of not knowing who you are? (COCA, fiction, 2003)
b. Even those of us who overtly scorn the idea of gentlemanliness may feel ashamed of not behaving like gentlemen (or ladies, for that matter). (COCA, academic, 2002)

The absolute frequency of the *of -ing* type is higher than that of the *to* infinitive, which naturally also entails that its proportion is much higher than the overall 15.2%. Based on this, it could be inferred that negation in the complement is

a feature that favors the *of -ing* pattern. In addition, because with the *to* infinitive the negator *not* can appear either before or after *to* (see e.g. Huddleston & Pullum 2002: 803), it might be difficult for speakers to decide which one is the "correct" position, and they opt for the *of -ing* complement, in which the only place for the negator is after the preposition, before the gerund. This would be a very practical explanation for the preference of the *of -ing* complement in these environments.

However, there are two further points to consider. The first one is the obvious problem with low frequency. With only nine instances altogether, nothing definite can be said. The second one is that four of the seven *of -ing* instances with negation in the complement clause are also non-agentive (see example 9a). This is a feature that has been observed to be associated with the gerund (Rudanko 2006, 2010, 2012, 2014, 2015), also in particular with *ashamed* (Rudanko 2017). Thus, it cannot be said with certainty whether in these cases the choice of *of -ing* is due to the negation or the non-agentive nature of the lower clause event.

Regarding the feature 'non-agentive', and also the feature 'passive' since it also renders the subject of the sentence non-agentive, it looks like the proportions of *of -ing*, 18.8% and 19.0%, are slightly higher than the average 15.2%, but only slightly. However, this indicates that the feature might have some effect on the choice of complement, especially because this is also noted in Rudanko's (2015, 2017) investigations in which he observed that the *of -ing* pattern more readily occurs in non-agentive than agentive environments with the adjectives *afraid* and *ashamed*, and the same trend was observed with the *to -ing* pattern by Rudanko (2012, 2015) regarding the adjectives *prone* and *accustomed*. Nevertheless, in the data set in the present study the numbers are too similar to the average 15.2% for any notable difference to be observed.

If we also think about the semantics of the feature 'copula' (see examples in (10)), we can observe that usually there is very little agentivity in the particular situation on the subject's part.

(10) a. And all because you're ashamed of being a Jew! (COCA, fiction, 2011)
b. You should not be ashamed of being kind. (COCA, fiction, 2007)
c. I am so ashamed of being in a bikini swimsuit. (COCA, magazine, 2002)

The non-agentive nature of copular expressions is of course noted in many works that discuss and define semantic roles, for instance Givón (2001), Huddleston & Pullum (2002), Payne (2006) and many others. This makes it an attractive option to analyze copular expressions as non-agentive, just like passives and non-agentive instances. As we can see in Table 2, there were 127 instances of copular expressions in the data, of which 76 had the *to* infinitive and 51 had the *of -ing* complement. Even though the *of -ing* option is in the minority, the proportion (40.2%) is much

higher than the average 15.2% for the *of -ing* complement, which indicates that the copular construction has an effect on the complement selection. Interestingly, a closer look at the types of copular expressions in the lower clauses reveals that especially adjectival predicatives tend to favor the *of -ing* complement: there were altogether 48 instances of copula + AdjP expressions, of which 28 (58.3%) had the *of -ing* complement (compare to copula + NP expressions, of which 15/55 (27.3%) had the gerund). The reason for this preference is not clear, but it may have something to do with the fact that adjectives more often than nouns denote (possibly) temporary states (*ashamed of being fat/sad/shy*), which is more in line with the ongoing but temporary nature of the gerund than the *to* infinitive. Anyway, this last point would need more careful investigation and is out of the scope of the present study.

All in all it seems that the three categories that have non-agentive properties draw the proportions from the average more towards the *of -ing* pattern, and while the passive and non-agentive numbers are close to the average proportion and no definitive conclusions can be drawn regarding them, copular expressions clearly favor the *of -ing* complement more than would be expected based on the average proportions.

The perfect aspect in the complement was also examined to see whether it has an effect on the complement choice. In Table 2 we can see that the proportion of *of -ing* in cases with the perfect aspect is about twice as high as the expected proportion. This would indicate that the perfect aspect somewhat attracts the *of -ing* complement. What is interesting though is that in seven cases out of the total 17 in which the perfect aspect occurs with the *to* infinitive, there is also another element present that has been observed to possibly attract the *of -ing* complement: in two cases a copula, in two cases a passive, and in three cases a non-agentive event. This means that in these seven cases, there are two features that in theory should attract the *of -ing* complement, but still they occur with the *to* infinitive. Further inspection of these seven instances did not shed more light on the matter, unfortunately. All in all, the results from the investigation of the effect of the perfect aspect are inconclusive.

If we examine the rest of the results in Table 2, negation in the main clause does not seem to have an effect on the complement choice as the numbers with that particular feature are almost similar to the overall proportions. Extraction contexts on the other hand seem to favor the *to* infinitive, as predicted by the Extraction Principle (Vosberg 2003), beyond the expected proportions.

Lastly, when examining the data, a recurring pattern was noted: *ashamed to say/admit*, or occasionally another verb of similar type, followed by a *that*-clause (overt or zero *that*). It seemed to be fairly frequent, so closer inspection

was deemed necessary. All in all, there were 160 instances of this type (84 with *say* and 61 with *admit*). Rudanko (2017: 52–53) also noted this pattern in COCA and COHA, and stated that it may be an emerging fixed expression. This could have been grounds to exclude it from an investigation such as the current one since there is no variation with the *of -ing* complement. These tokens were not, however, excluded from the relevant cases, since there is at least some variation: in the data, three instances were found with the *of -ing* complement, exemplified in (11).

(11) a. I'm not ashamed of saying I'm there to sell you something (COCA, spoken, 1990)

b. …nor am I ashamed of believing that OJ got away with murder. (COCA, news, 1995)

c. …and are not ashamed of proclaiming that it is God who gives us our rights. (COCA, spoken, 2012)

Two of the examples are from the spoken part of COCA, and the example from the news part seems to be taken from an opinion segment in the *San Francisco Chronicle*, which includes spoken-like text, which might explain that the otherwise more common *to* infinitive pattern was not used. However, based on the data it can be said that the pattern *ashamed to say/admit* + *that*-clause, with those two specific verbs, is a fairly well established construction which could be worth investigating further.

5. Summary and conclusion

The present paper set out to investigate the adjective *ashamed* and the variation between the two complement types that occur with it: *to* infinitive and *of -ing*. The first aim was to study the diachronic development of *ashamed* and its different complements, and the second was to analyze which features affect the choice of the complementation pattern.

The adjective *ashamed* has lost ground during the 200-year period recorded in COHA. At first in the 19th century there was a period of increase, but after the 1870s the frequency began to decrease from 37 instances pmw to 12 instances pmw in the 2000s. The same decreasing trend in frequency can be seen in *ashamed* with the two complement types. Regarding the development of all the different patterns in which *ashamed* occurs, the most notable change has happened in the proportion of the *to* infinitive: in the 1830s it covered almost half of the occurrences of *ashamed*, but in the 2000s only about 20%. The *of*

-ing pattern has been fairly infrequent in all decades, so it is not the pattern that has replaced the *to* infinitive. Instead, the complementless use of *ashamed* has increased steadily since the 19th century, and in the 2000s data *ashamed* usually occurs without a complement, or with an *of* + NP complement that has been frequent throughout the past two centuries. The reasons for the increased preference of not having a complement clause are unclear and would need further research.

The investigation into the features that might have an effect on the complement choice showed that while the *to* infinitive is overall much more frequently used than *of -ing*, there is one exception: when the complement is negated, the *of -ing* pattern is favored. However, the token count for these cases was quite low, and there was overlap with non-agentive cases, so more data would be needed to confirm this tendency. In addition, it was observed that the *of -ing* pattern is present more often than expected when the complement has the perfect aspect or copular construction. With the perfect aspect the token counts were fairly low, but with the copular constructions there were 127 instances and the proportion of *of -ing* in these contexts was clearly higher than the average (40.2% vs. 15.2%), which indicates that copula in the complement attracts the *of -ing* pattern. Given that copular constructions can be seen as non-agentive, or [–Choice], environments, this would provide some further evidence for Rudanko's (2017) Choice Principle. The investigation also found confirmation for the Extraction Principle, as the *to* infinitive was chosen in 93% of the extraction environments. Finally, it was noted that the pattern *ashamed to say/admit* + *that*-clause is quite frequent in the data, and since there is very little variation with the *of -ing* pattern, it may be considered a fixed phrase.

All in all it is interesting that the complementless use of *ashamed* has increased at the expense of the *to* infinitive complement, and a further step would be to investigate this in detail in order to find the reasons behind this, and also to see whether this is something that has happened with other adjectives. Regarding the different features possibly affecting the complement choice, the analysis of the perfect aspect and negation in the complement would definitely benefit from more data, possibly in the form of larger corpora or expansion of investigation to other adjectives.

Acknowledgements

I want to thank my co-editors and the two anonymous reviewers for their very helpful and insightful comments that prompted major changes in the present chapter. All the remaining errors and shortcomings are my own.

References

Primary sources

Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 520 million words, 1990–2015*. <http://corpus.byu.edu/coca/>

Davies, Mark. 2010–. *The Corpus of Historical American English: 400 million words, 1810–2009*. <http://corpus.byu.edu/coha/>

Secondary sources

Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2: 119–127.

De Smet, Hendrik. 2013. *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: OUP.

Egan, Thomas. 2008. *Non-finite Complementation. A Usage-based Study of Infinitive and -ing Clauses in English*. Amsterdam: Rodopi.

Fanego, Teresa. 1997. On patterns of complementation with verbs of effort. *English Studies* 78: 60–67. doi:10.1080/00138389708599061

Givón, Talmy. 1985. Iconicity, isomorphism and non-arbitrary coding in syntax. In *Iconicity in Syntax* [Typological Studies in Language 6], John Haiman (ed.), 187–219. Amsterdam: John Benjamins. doi:10.1075/tsl.6.10giv

Givón, Talmy. 1991. Isomorphism in the grammatical code: Cognitive and biological considerations. *Studies in Language* 15(1): 85–114. doi:10.1075/sl.15.1.04giv

Givón, Talmy. 2001. *Syntax*, Vol. I & II. Amsterdam: John Benjamins.

Havu, Jukka & Höglund, Mikko. 2015. A cross-linguistic perspective on complementation in the tough construction. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 52–74. Houndmills: Palgrave Macmillan.

Herbst, Thomas, Heath, David, Roe, Ian F. & Götz, Dieter. 2004. *A Valency Dictionary of English. A Corpus-Based Analysis of the Complementation Patterns of English Verbs, Nouns and Adjectives*. Berlin: Mouton de Gruyter. doi:10.1515/9783110892581

Huddleston, Rodney & Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Höglund, Mikko. 2014. Active and passive infinitive, ambiguity and non-canonical subject with ready. In *Corpus Interrogation and Grammatical Patterns* [Studies in Corpus Linguistics 63], Kristin Davidse, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds), 239–262. Amsterdam: John Benjamins.

Kjellmer, Göran. 1980. Accustomed to swim; accustomed to swimming: On verbal forms after TO. In *ALVAR: A Linguistically Varied Assortment of Readings. Studies Presented to Alvar Ellegård on the Occasion of his 60th Birthday* [Stockholm Papers in English Language and Literature 1], Jens Allwood & Magnus Ljung (eds), 75–99. Stockholm: University of Stockholm, Department of English.

Mair, Christian. 2006. *Twentieth-Century English. History, Variation and Standardization*. Cambridge: CUP. doi:10.1017/CBO9780511486951

Mindt, Ilka. 2011. *Adjective Complementation: An Empirical Analysis of Adjectives Followed by that-Clauses* [Studies in Corpus Linguistics 42]. Amsterdam: John Benjamins. doi:10.1075/scl.42

OED = *Oxford English Dictionary*, online edn. (May 2017).
Payne, Thomas E. 2006. *Exploring Language Structure*. Cambridge: CUP. doi:10.1017/CBO9780511806483
Rohdenburg, Günter. 1995. On the replacement of finite complement clauses by infinitives in English. *English Studies* 4: 367–388. doi:10.1080/00138389508598980
Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2): 149–182. doi:10.1515/cogl.1996.7.2.149
Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms: English from 1500–2000*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166. Bern: Peter Lang.
Rohdenburg, Günter. 2015. The embedded negation constraint and the choice between more or less explicit clausal structures in English. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 101–127. Houndmills: Palgrave Macmillan.
Rudanko, Juhani. 1999. *Diachronic Studies of English Complementation Patterns. Eighteenth Century Evidence in Tracing the Development of Verbs and Adjectives Selecting Prepositions and Complement Clauses*. Lanham MD: University Press of America.
Rudanko, Juhani. 2000. *Corpora and Complementation. Tracing Sentential Complementation Patterns of Nouns, Adjectives and Verbs over the Last Three Centuries*. Lanham MD: University Press of America.
Rudanko, Juhani. 2006. Watching English grammar change: A case study on complement selection in British and American English. *English Language and Linguistics* 10(1): 31–48. doi:10.1017/S1360674306001791
Rudanko, Juhani. 2010. Explaining grammatical variation and change: A case study of complementation in American English over three decades. *Journal of English Linguistics* 38(1): 4–24. doi:10.1177/0075424209335495
Rudanko, Juhani. 2011. *Changes in Complementation in British and American English: Corpus-Based Studies on Non-Finite Complements in Recent English*. Houndmills: Palgrave Macmillan. doi:10.1057/9780230305199
Rudanko, Juhani. 2012. Exploring aspects of the Great Complement Shift, with evidence from the TIME Corpus and COCA. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Closs Traugott (eds), 222–232. Oxford: OUP.
Rudanko, Juhani. 2014. A new angle on infinitival and *of -ing* complements of *afraid*, with evidence from the TIME Corpus. In *Corpus Interrogation and Grammatical Patterns* [Studies in Corpus Linguistics 63], Kristin Davidse, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds), 223–238. Amsterdam: John Benjamins.
Rudanko, Juhani. 2015. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Houndmills: Palgrave Macmillan.
Rudanko, Juhani. 2017. *Infinitives and Gerunds in Recent English: Studies on Non-Finite Complements with Data from Large Corpora*. London: Palgrave Macmillan.
Van linden, An. 2010. The rise of the *to*-infinitive: Evidence from adjectival complementation. *English Language and Linguistics* 4(1): 19–51. doi:10.1017/S1360674309990396
Van linden, An. 2012. *Modal Adjectives: English Deontic and Evaluative Constructions in Diachrony and Synchrony*. Berlin: Mouton de Gruyter. doi:10.1515/9783110252941

Van linden, An & Davidse, Kristin. 2009. The clausal complementation of deontic-evaluative adjectives in extraposition constructions: A synchronic-diachronic approach. *Folia Linguistica* 43(1): 175–217.

Vosberg, Uwe. 2003. The role of extractions and *horror aequi* in the evolution of *-ing* complements in Modern English. In *Determinants of Grammatical Variation in English* [Topics in English Linguistics 43], Günter Rohdenburg & Britta Mondorf (eds), 305–327. Berlin: Mouton de Gruyter. doi: 10.1515/9783110900019.305

Vosberg, Uwe. 2006. *Die Große Komplementverschiebung: Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.

Sentential complementation of *propose* in recent British English

Veera Saarimäki
University of Tampere

This article examines the sentential complementation of *propose* from 1780 to the 1990s. The study is based on previous research on the use of the verb in the Corpus of Late Modern English Texts and the British National Corpus. The focus of the article is on control theory, and it tracks changes in the complement patterns and control structure of *propose*. The examination of the corpora shows that throughout the data, the *to*-infinitive is the most frequent complement of *propose*, followed by the *that*-clause and *-ing* clause. The *to*-infinitive has increased its proportion at the expense of other patterns. The control structure of *propose* is complex, as it allows subject, object, PP object as well as unspecified control, the latter increasing in use. The data indicate that most changes occur between 1780 and 1920, reinforcing the idea that the Late Modern English period was a time of significant change in the complementation systems of many predicates.

Keywords: English; corpus linguistics; complementation; control; propose

1. Introduction

In the field of modern English linguistics it is generally acknowledged that, as heads of phrasal constructions, verbs have an effect on some of the items that follow them. Consider the following examples drawn from the Corpus of Late Modern English Texts (CLMET):

(1) Bellarius also *proposed* to carry her out into the forest, and there celebrate her funeral with songs and solemn dirges, as was then the custom.
(1807 Lamb, *Tales from Shakespeare*)

(2) He *proposed* that if the man returned he should not be interfered with, but merely watched from the other side of the glass door.
(1902 Bennett, *The Grand Babylon Hotel*)

(3) The young woman whom I mentioned to you *proposed* rowing me across the water amongst the rocks.
(1796 Wollstonecraft, *Letters on Sweden, Norway and Denmark*)

DOI 10.1075/slcs.195.07saa

Sentences 1–3 exemplify the most frequently used sentential complements of the verb *propose*, i.e. the *to*-infinitive, *that*-clause and *-ing* clause. This article tracks the development of these three as well as other, less common, sentential complement patterns of *propose* from 1780 to the end of the twentieth century with the help of two corpora, the CLMET and the British National Corpus (BNC). Diachronic surveys in the complementation of this particular verb have not been previously conducted with this set of corpus data, so the aim is to contribute to the scholarly treatment of verb complementation with this focused study.

The article is based on my MA thesis (Saarimäki 2014), which studied both the sentential and non-sentential complementation of *propose* extensively, drawing on theories such as theta theory (e.g. Haegeman 1991: 41ff), the extraction principle (e.g. Vosberg 2003: 308), and control theory (e.g. Davies & Dubinsky 2004). In the current work the theoretical framework has been limited to control theory, as the control structure of *propose* provided some of the most interesting findings in the MA thesis.

The paper starts by looking at *propose* in the literature to form an initial idea of the sentential complements that can be expected to occur with the verb in the corpus data, as well as the control tendencies of the verb (Section 2). Section 3 introduces the data and methods, and Section 4 discusses the findings. Conclusions are presented in Section 5.

2. *Propose* in the literature

This section looks at the sentential complementation of *propose* on the basis of well-known dictionaries and grammars, introduces control theory, and summarises any remarks on the control structure of *propose* found in the dictionaries and grammars. The findings act as a starting point for the analysis of the authentic corpus data.

2.1 *Propose* in dictionaries

In order to see what kinds of complementation patterns the verb *propose* has been presented as selecting, two dictionaries that differ in both style and purpose were consulted: the *Oxford English Dictionary* (*OED*), which focuses on listing, describing and exemplifying the senses of words, and Poutsma's (MS) unpublished *Dictionary of Constructions*, which focuses on the different constructions in which words occur. Both dictionaries provide multiple senses for the verb, two of which are most likely to occur with sentential complementation and thus the most relevant for the current paper. These two senses, as defined in the *OED* (s.v. *propose*), are given in Table 1.

Table 1. The relevant senses of *propose* (*OED*, s.v. *propose*)

1. With infinitive or verbal noun as object. To set before oneself as something to be done; to intend, purpose, or design.
2. With clause, infinitive, or simple object. To put forward or suggest as a scheme, plan, or course of action; to recommend or advocate that something be done.

Moving on to the syntax of *propose*, Table 2 summarises the sentential complement patterns identified in the examples provided by the dictionaries.

Table 2. Sentential complement patterns of *propose* in the dictionaries

OED	**Poutsma**
to-infinitive	*to*-infinitive
NP + *to*-infinitive	NP + *to*-infinitive
to + NP + *to*-infinitive	*to* + NP + *to*-infinitive
that-clause	*that*-clause
-*ing* clause	-*ing* clause
	poss. + -*ing* clause

As the table shows, *propose* occurred with a total of six patterns in the dictionaries, five of which were exemplified in both. Of the five, the NP + *to*-infinitive pattern stands out; Poutsma (MS, s.v. *propose*), in fact, notes that it is "an unusual one". Indeed, the pattern is only exemplified once in each dictionary:

(4) Þei..may *propose* hem to gete or noye þe townes. (*OED*: 1441 H. Nicolas)

(5) The monopoly plan . . is *proposed* to be applied to sago and sugar. (Poutsma: *Daily Mail*)

The four other complement patterns seem to be more common and can be expected to be found in both the CLMET and the BNC, whereas the poss. + -*ing* clause pattern might be rarer, as indicated by its absence from the *OED*:

(6) You *proposed* our confessing our faults. (Poutsma: Hardy, *Tess*)

2.2 Propose *in grammars*

For further information on the syntax of *propose*, three influential grammars were considered. A summary of the sentential complement patterns given for *propose* in the grammars is presented in Table 3.

Table 3. Sentential complement patterns of *propose* in the three grammars

Quirk et al. (1985) ***A Comprehensive Grammar of the English Language***	**Biber et al. (1999)** ***Longman Grammar of Spoken and Written English***	**Huddleston and Pullum (2002)** ***Cambridge Grammar of the English Language***
to-infinitive	*that*-clause	*to*-infinitive
NP + *to*-infinitive	*to* + NP + *that*-clause	*for* + NP + *to*-infinitive
that-clause	*-ing* clause	*that*-clause
to + NP + *that*-clause	NP + *-ing* clause	*to* + NP + *that*-clause
-ing clause		*-ing* clause

The representation of *propose* in the grammars is relatively similar to the dictionaries in that the most frequently exemplified patterns, *to*-infinitive, *that*-clause and *-ing* clause, can be found in both Tables 2 and 3. However, there is more variation between the grammars than there is between the *OED* and Poutsma. Quirk et al. (1985: 1181–1183) list the NP + *to*-infinitive pattern that Poutsma (MS, s.v. *propose*) considered "unusual", whereas Huddleston and Pullum (2002: 1226) indicate that *propose* can also take a *for* + NP + *to*-infinitive complement. Biber et al. (1999) do not mention the *to*-infinitive in their list of complements of *propose*, whereas all three grammars state that the *that*-clause complement can be preceded by *to* + NP, a pattern which did not occur in either the *OED* or Poutsma.

2.3 Control theory and *propose*

According to Huddleston and Pullum (2002: 1194), a distinction can be made between two types of verbs:

(7) Liz hoped to convince them.

(8) Liz seemed to convince them.

In (7), *Liz* is an argument of *hope*, whereas *Liz* is not an argument of *seem* in (8). In other words, *Liz* is doing the hoping in (7), but it is not possible for her to be "seeming" in (8). This distinction is caused by the fact that *hope* is a *control predicate* and *seem* is an *NP movement predicate*. As *propose* functions similarly to *hope*, the concept of control requires further discussion.

In control theory, it is argued that the predicate in the complement clause has a subject, although it is often not structurally present. The argument is supported by theta theory: according to the theta criterion, "[e]ach argument is assigned one and only one theta role [and] [e]ach theta role is assigned to one and only one argument" (Haegeman 1991: 46). The first half of the criterion is relevant here, as Huddleston and Pullum (2002: 1194) point out that in (7) *Liz* seems to be assigned

two theta roles: Experiencer by *hope* and Agent by *convince*. Since the theta criterion forbids this, *convince* must have an understood subject, PRO, that is coreferential with, i.e. controlled by, the subject of the matrix clause, *Liz*:

(7′) Liz hoped [PRO to convince them.]

In (7′) the problem has been solved, as the theta role Agent can be assigned to the understood subject PRO.

The controlling element is not always the subject of the matrix clause, as exemplified in (9), adapted from Huddleston and Pullum (2002: 1193):

(9) Pat persuaded Kim [PRO to travel by bus].

In (9), PRO is controlled by, and is thus coreferential with, the object of *persuade*: *Kim*.

To start defining the control structure of *propose*, the literature discussed above was reviewed. According to Quirk et al. (1985: 1187), *propose* is always a subject control verb, when it is followed by a simple *to*-infinitive[1] or an *-ing* clause. They make no remarks on control in relation to any other complement patterns of *propose*, nor do the other grammars.

Poutsma (MS, s.v. *propose*), on the other hand, comments on the infinitival complement patterns of *propose* as follows: "…the subject of the action indicated by the infinitive may be the speaker(s), the person(s) spoken to, or the speaker and the person(s) spoken to together." The first of these refers to subject control, whereas in the second type the NP in the *to* + NP + *to*-infinitive or NP + *to*-infinitive patterns controls the subject of the *to*-infinitive.[2] This means that *propose* is both a subject and an object control verb, although it has to be noted that in object control the controlling element is traditionally an NP directly following the matrix verb, rather than an NP in a PP. To differentiate between control in the *to* + NP + *to*-infinitive and NP + *to*-infinitive patterns, I will call the former *PP object control*. The final type, henceforth *unspecified control*, refers to a kind of combination of subject and object control, in which the controller of the understood subject is not explicitly present, as in Poutsma's (MS, s.v. *propose*) example:

(10) He *proposed* to go into the first public-house we should find open.
(Poutsma: Smol., *Rod. Rand*)

1. As opposed to the *to* + NP + *to*-infinitive and NP + *to*-infinitive patterns.

2. I interpret Poutsma's (MS, s.v. *propose*) second type of control to refer to (PP) object control constructions in which the controlling element is explicitly present, and decided to include all instances without an explicit controller in the third type, although the understood subject might in some cases be "the person(s) spoken to" rather than "the speaker and the person(s) spoken to together".

On the basis of Poutsma's analysis, the subject of *go* in (10) is both *he* and the addressee(s), which suggests that the sentence could be rephrased as (10′):

(10′) He *proposed* that we should go into the first public-house we should find open.

However, it could be argued that *he* could be making the proposition to the hearers (*we*) alone without any intention of going to the public-house himself. In this case the *to* + NP element could be considered to have been dropped, and (10) could be rephrased as (10″):

(10″) He *proposed* to us to go into the first public-house we should find open.

On the basis of example (10), it seems to be possible to omit the controlling item, the PP preceding the *to*-infinitive, from the sentence. This is an interesting quality of *propose*, as Rudanko (1989: 138), discussing similar constructions and their underlying structures, notes that with verbs that occur in the PP object control pattern "the PP is in general much less freely omissible" than with subject control constructions. The statement is closely related to Bach's generalisation (based on the works of Bach 1979; 1980), which has been formulated by Rizzi (1986: 503) as follows: "In object control structures the object NP must be structurally represented." If the generalisation is thought to include prepositional objects, *propose* violates it.

To summarise, the control structure of *propose* is diverse. It will be interesting to see whether examples of all four types of control can be found in the authentic data, and whether changes have occurred in their use over time.

3. Data and methods

The historical data for the study were collected from the Corpus of Late Modern English Texts (CLMET). The second (1780–1850; 3.7 million words) and third (1850–1920; 4 million words) periods of the CLMET were searched for all forms of *propose*, i.e. *propose*, *proposes*, *proposed* and *proposing*, the corpus being untagged at the time. The smaller, original version of the CLMET was chosen over the more recently updated extended version and version 3.0 of the corpus, as it was large enough to provide an appropriate number of tokens for the study.

The more recent data, containing material from the 1960s to 1990s, were drawn from the British National Corpus (BNC) with a lemma search. The search was limited to the imaginative prose domain (16.5 million words) of the BNC to keep the number of hits reasonable and to ensure that the data would be stylistically similar

to the CLMET data. It should be noted, however, that the CLMET is more varied in its text types than the imaginative prose domain of the BNC, which could have an effect on the findings.

After the data were collected, the tokens were examined manually to identify the complement of *propose* in each instance. In cases where distinguishing between a complement and an adjunct was problematic, I relied on Somers's (1984: 516–517) *do so* test, according to which the "*do so* phrase can be the proform of anything up to the entire predication […], [but] the minimum element that can be substituted is the predicate plus any complements" (1984: 517). Thus, (12) below is rendered ungrammatical by the complement of *propose*, "a toast", not being included in the proform *do so*.

(11) Jake *proposed* a toast at the birthday party and Mary *did so* at the graduation.

(12) *Jake *proposed* a toast at the birthday party and Mary *did so* a toast at the graduation.

In addition to identifying the complements, irrelevant tokens were discarded. There were a total of 12 irrelevant tokens in the second period of the CLMET (CLMET2), 14 in the third (CLMET3) and one in the BNC. All of these were examples of the adjectival use of *proposed*. After the manual pruning, the number of tokens to be analysed is 298 for the CLMET2, 377 for the CLMET3, and 368 for the BNC. Because of the differences in the total word counts of the subcorpora, raw counts are normalised to frequencies per 1,000,000 words in the discussion below.

4. Findings and discussion

This section discusses developments in the sentential complementation and control structure of *propose* from 1780 to the end of the twentieth century.

4.1 Sentential complementation of *propose*

4.1.1 *Overview*

The total number of tokens in the subcorpora is 1,043, divided rather equally between the three periods: 298 in the CLMET2, 377 in the CLMET3 and 368 in the BNC. However, if these numbers are compared to the total word counts of the subcorpora, 3.7, 4 and 16.5 million words respectively, it is quite apparent that the overall use of *propose* has decreased notably in the twentieth century. Figure 1 illustrates the change.

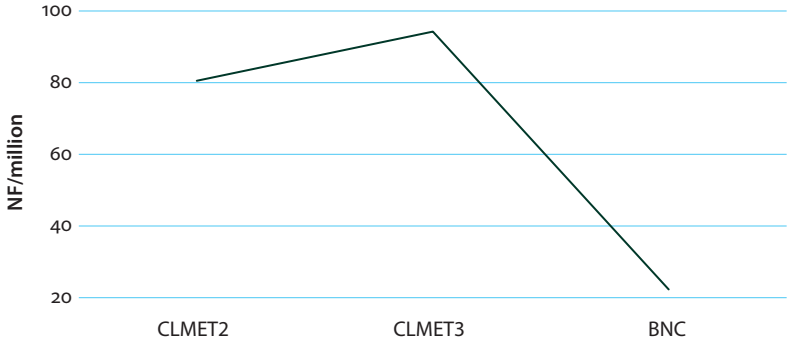

Figure 1. Changes in the frequency of *propose* from 1780 to the 1990s

In the period 1780–1850, the frequency of the verb *propose* is 80.5 per million words, after which it rises to 94.3 in the period 1850–1920. During the approximately 40-year gap in the data, the number drops to roughly one fourth, i.e. 22.3 occurrences per million words. One reason for the sharp decline might be that the CLMET contains treatises and other non-fictional text types that favour more formal verbs, whereas the imaginative prose domain of the BNC consists entirely of fictional texts. However, the difference in style is unlikely to be the only reason for the drop, as the earlier data does contain plenty of tokens from fictional material.

Between the three subcorpora, *propose* occurs with nine different sentential complement patterns, although no one subcorpus contained all nine complement types. Three patterns are variations of the *to*-infinitive, and there are two variations of both *that*-clauses and *-ing* clauses in the data. The two remaining complement patterns are the quote and the *wh*-clause, both of which were only found in two of the subcorpora. The patterns will be discussed in four sections: *to*-infinitive patterns, *-ing* clause patterns, *that*-clause patterns, and other patterns.

Because of the decrease in the overall use of *propose*, it is likely that most complement patterns will have gone through a similar change. For this reason, relying on normalised frequencies (NFs) alone would not provide a comprehensive picture of the developments that have taken place in the complement structure of *propose*. In addition to the NFs, the proportion of each complement of the total number of tokens will be taken into account.

4.1.2 To-*infinitive patterns*

Throughout the data, the *to*-infinitive is the most common sentential complement of *propose*, and after the CLMET2 period, it is the most common complement of *propose* overall:

(13) When Aunt Sarah arrived, *proposing* to take them both to live with her in Coniston, he had utterly refused to go. (BNC: F99 36)

The pattern has a total of 75 tokens in the CLMET2, 163 in the CLMET3 and 128 in the BNC. Figure 2 shows the NF of the *to*-infinitive as well as its proportion in each subcorpus.

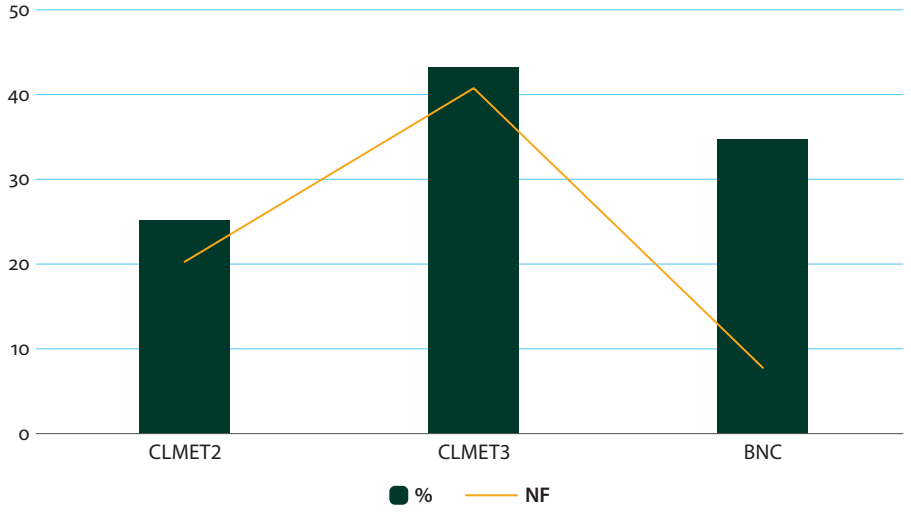

Figure 2. Normalised frequency and proportion of the *to*-infinitive complement in the subcorpora

As Figure 2 shows, both the NF and the proportion of the *to*-infinitive as much as double from the CLMET2 to the CLMET3, after which the NF drops to less than one fourth in the BNC, following the overall frequency of the verb. The proportion of the complement pattern does not decrease as steeply, but instead settles to the middle ground between the two earlier periods. Closer inspection reveals that the steep increase of the complement in the CLMET3 can be pinpointed to one text: out of the 163 *to*-infinitives, as many as 57 occur in one text, namely Booth's *In Darkest England and the Way out*:

(14) I *propose* to multiply their number, to develop their usefulness, and to make them the threshold of the whole Scheme. (CLMET3: 1890 Booth, *In Darkest England and the Way out*)

If this text is removed from the calculations, the *to*-infinitive has a NF of 27.2 and a proportion of 35.3% of all complements of *propose* in the CLMET3. Both numbers point to an increase in the use of the pattern in the period 1850–1920, after which the situation has stabilised, the *to*-infinitive accounting for 34.8% of the tokens in the BNC.

The two other *to*-infinitive patterns found in the data are *to* + NP + *to*-infinitive and NP + *to*-infinitive. The former occurs seven times in both the CLMET2 and

the CLMET3, which indicates a slight decrease in both NF and proportion before the pattern, according to the BNC data, falls out of use:

(15) It is just what I *proposed* to her to do–to have crossed the Alps with me, to sail on sunny seas… (CLMET2: 1823 Hazlitt, *Liber Amoris*)

Considering that none of the grammars listed the pattern as a possible complement of *propose*, the development is not surprising.

The NP + *to*-infinitive complement is rare in all subcorpora, with just two tokens in the CLMET2, five in the CLMET3 and one in the BNC. Judging by the few available instances of the pattern, it tends to occur in passivized sentences and has a formal feel about it:

(16) Apart from the independent agencies employed to prosecute this class on enquiries, which it is *proposed* to very largely increase, the Army possesses in itself peculiar advantages for this kind of investigation (CLMET3: 1890 Booth, *In Darkest England and the Way out*)

On the basis of the corpus data, the inclusion of the pattern in Quirk et al. (1985: 1181–1183) as a complement of *propose* feels trivial considering that other rare patterns are not mentioned.

4.1.3 -ing *clause patterns*

The *-ing* clause occurs as a complement of *propose* a total of 23 times in the CLMET2, 12 in the CLMET3 and 11 in the BNC. Figure 3 shows how the normalised frequency and proportion of the pattern have decreased over time.

The *-ing* clause is often regarded as a rival of the *to*-infinitive as a result of some of the developments in the Great Complement Shift (see e.g. Rohdenburg 2006; Rudanko 2012; Vosberg 2009). However, *propose* appears to be a counterexample to the theory, as the use of the *to*-infinitive as a complement of *propose* has increased at the expense of the *-ing* clause (and other sentential complements). According to Vosberg (2009: 217), after the *-ing* form has established itself as a complement of a particular verb, it can have been "felt to be semantically inadequate", which has resulted in a decline. Although the current data do not show how the *-ing* clause gained ground as a complement of *propose*, the decline was clearly on-going between 1780 and 1920.[3]

3. See Fanego (2007: 182–186) for a discussion on *propose* and the development of the *-ing* clause complement. In her 18th-century data *-ing* clause complements are recorded with *propose* in sense 2 "in ten cases, *to*-infinitives only in five."

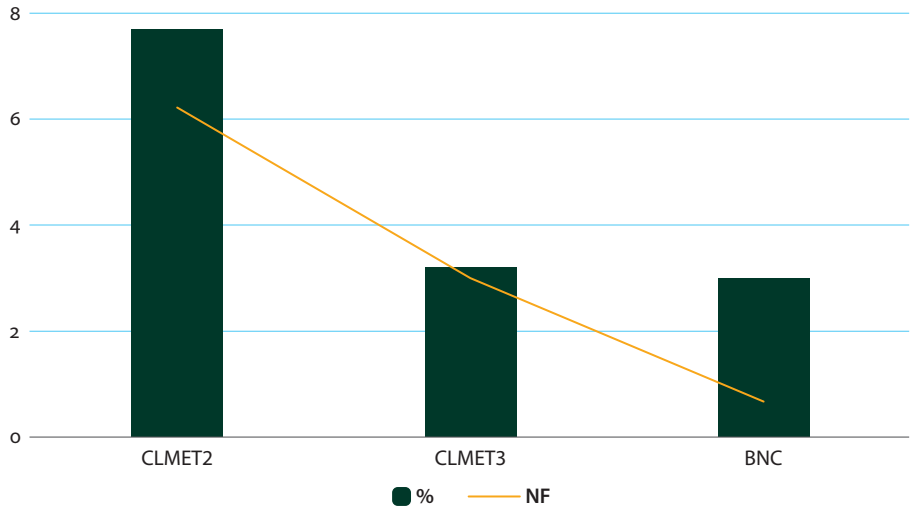

Figure 3. Normalised frequency and proportion of the *-ing* clause complement in the subcorpora

The other *-ing* clause pattern in the data is the possessive + *-ing* clause complement, which was only found in three tokens, all from the CLMET2:

(17) In appointing a morning's or an evening's walk, he *proposed* HER [sic] going with the rest; no one had ever required her company before.
(CLMET2: 1796 Inchbald, *Nature and Art*)

The pattern was mentioned in both Poutsma (MS, s.v. *propose*) and Biber et al. (1999: 742), but on the basis of the current data it seems to have disappeared from use or become extremely rare.

4.1.4 That-*clause patterns*

The second most common sentential complement of *propose* in all three subcorpora is the *that*-clause. It can follow *propose* on its own or the recipient of the proposal can be inserted between the two items in the form of *to* + NP:

(18) I *propose* that it should be treated as a case of criminal kidnap until we have evidence to the contrary. (BNC: CEC 3169)

(19) In consequence of this reasoning, Cecilia formed a design of *proposing* to her companions that they should give a prize…
(CLMET2: 1796–1801 Edgeworth, *The Parent's Assistant*)

The latter variant of the complement is surprisingly infrequent, considering it was mentioned in all three grammars. The *to* + NP + *that*-clause pattern has seven

tokens in the CLMET2 and two in both the CLMET3 and the BNC: even in the CLMET2 it only accounts for 2.3% of all complements of *propose*.

The straightforward *that*-clause, on the other hand, is clearly more common with 53, 37 and 35 hits. The NFs and proportions for each subcorpus are presented in Figure 4.

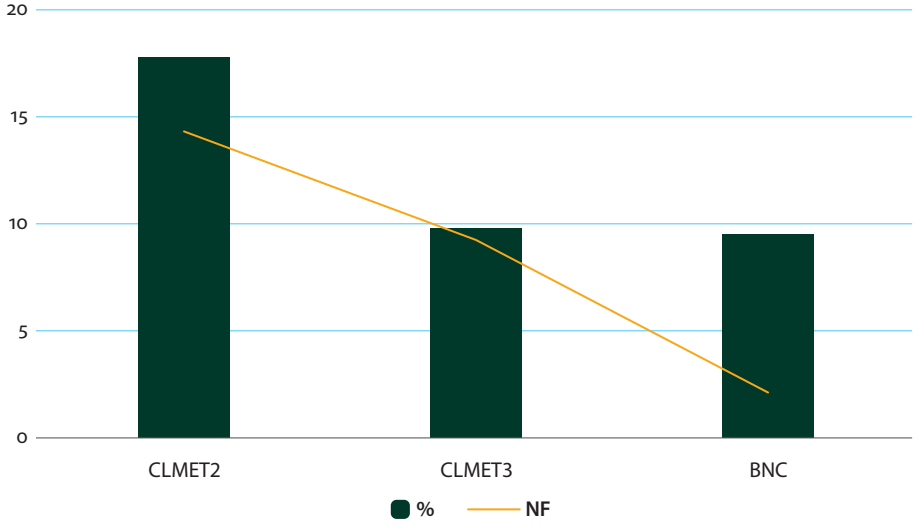

Figure 4. Normalised frequency and proportion of the *that*-clause complement in the subcorpora

As the decrease in the proportions of both the *that*-clause and the *-ing* clause takes place at the same time as the increase in the proportion of the *to*-infinitive does, it is likely that the latter has gained ground at the expense of both of them. In general, the Great Complement Shift contrasts the *to*-infinitive with the *-ing* clause, but as Rohdenburg (2006: 143) points out, the *that*-clause was also affected.

4.1.5 *Other patterns*

Two other sentential complement patterns were also identified in the data, namely the quote and the *wh*-clause. The former has three tokens in the CLMET2 and four in the BNC, but none in the CLMET3:

(20) Mr. Towlinson *proposes* with a sigh, 'Amendment to us all'…
(CLMET2: 1848 Dickens, *Dombey and Son*)

The latter was not found in the earliest data, and it had only one hit in both the CLMET3 and the BNC:

(21) He cannot go down to Congress himself and *propose* what he wants: he can only write a letter and send it.
(CLMET3: 1867 Bagehott, *The English Constitution*)

Both of these patterns are rare with *propose*, and were not mentioned in the literature reviewed in Section 2.

4.2 Control

As was established in Section 2.3, *propose* has an interesting and somewhat complicated control structure. Subject control is the dominant control type in all three subcorpora with both *to*-infinitive and *-ing* clause complements:

(22) …and he *proposes* [he]$_{PRO}$ to hire a carriage for a whole month.
(CLMET2: 1821 Galt, *Ayrshire Legatees*)

(23) I *propose* [I]$_{PRO}$ giving you the first call, my old friend the second, and Miss Nimmo as I return home.
(CLMET2: 1780–96 Burns, *The Letters of Robert Burns*)

On the basis of Poutsma's remarks, some instances of object control and PP object control were also expected to be found. Both types are rare, as they are tied to the complement patterns NP + *to*-infinitive and *to* + NP + *to*-infinitive. All eight tokens of the former are examples of object control, and all 14 of the latter examples of PP object control:

(24) …and he asked me what I meant to do with it; *proposed* a Charity [a Charity]$_{PRO}$ to be established on behalf of decayed half-castes…
(CLMET3: 1870 Meredith, *The Adventures of Harry Richmond*)

(25) [H]er suspicion had then been first excited by his displeasure at her *proposing* to him [he] $_{PRO}$ to return it…
(CLMET3: 1865 Yonge, *The Clever Woman of the Family*)

Although the analysis presented in (25) illustrates how the NP *him* in the PP *to him* controls the subject of *return*, it is worth considering whether the PP is actually an element of the matrix clause or the lower clause. Rudanko (1989: 133–142) discusses similar constructions and their underlying structures. The verbs that he examines in more detail are *rely on*, *depend on* and *count on*. On the basis of five tests he (1989: 137) concludes that these particular verbs allow for both structures: the NP of the PP can be a part of either the higher or the lower clause, but remarks that there are verbs that only allow the interpretation whereby the NP belongs to the higher clause and an understood subject, PRO, is identified in the lower clause.

According to the tests, some of which are illustrated in (26–28), *propose* seems to be of this latter type:[4]

(26) *What this fellow *proposed* to was for the Sheik to put me to death.

(27) Who did this fellow *propose* to to put me to death?

(28) The person who this fellow was *proposing* to to put me to death was the Sheik.

In fact, Rudanko (1989: 140–142) refers to Visser's (1973: 2241ff.) list of verbs occurring in this particular pattern, and *propose* is one of the verbs included. However, according to Rudanko (1989: 142), the use of *propose* in the *to* + NP + *to*-infinitive pattern is "[v]ery marginal", which seems to be a correct assessment, considering that *propose* did not occur with this complement in the most recent data.

The final type of control listed by Poutsma (MS, s.v. *propose*) is unspecified control, in which the controlling element is not present in the sentence. In the CLMET2, there are at least[5] four instances of this control type with *to*-infinitives and two with *-ing* clauses:

(29) I *propose* to cut all the timber we want for the houses out of this part of the grove… (CLMET2: 1841 Marryat, *Masterman Ready*)

(30) He *proposed* sending delegates to entreat the assistance of other Trades Unions in other towns. (CLMET2: 1848 Gaskell, *Mary Barton*)

If we recall the method introduced in Section 2.3 whereby non-finite clauses are converted to finite *that*-clauses, the sentence in (29) might begin with "I *propose* that we cut…" and (30) with "He *proposed* that we/they send…" In (29) the preference for control involving both the speaker and the hearer is made stronger by the personal pronoun *we* functioning as the subject of *want* in the subordinate clause.

Duffley (2000: 221) readily recognises the possibility of this unspecified control with the *-ing* clause complement. He (ibid: 237) notes that the subject of the *-ing* clause complement is defined by the semantics of the matrix verb, "the

4. Examples (26)–(28) are modified versions of the following token from the CLMET2: "My donkey-boys afterwards said they had overhead [sic] this fellow *propose* to the Sheik to put me to death…" (CLMET2: 1844 Kinglake, *Eothen*).

5. I say *at least* because in some cases even a larger context does not reveal whether subject or unspecified control was intended.

grammatical meaning of the *-ing*, and the function of the *-ing* with respect to the matrix". He gives an example with *propose*: "He proposed seeing a psychiatrist.", and states that the action expressed by the *-ing* clause is "logically understood" (ibid: 238) as not being done by the person proposing it. Duffley also recognises the possibility of a sentence allowing a reading of both subject and unspecified control, especially with limited context.

Duffley's approach to the *to*-infinitive, on the other hand, is slightly different. He argues that "when the *to*-infinitive is used as a complement of another verb, it always has the same 'subject' as the matrix" (ibid: 235), which does not seem to be in agreement with the data presented in the current study and Poutsma's (MS, s.v. *propose*) observation of "the subject of the action indicated by the infinitive" being "the speaker(s), the person(s) spoken to, or the speaker and the person(s) spoken to together". Duffley (ibid: 240) gives an example, "John said to be careful", which, as he points out, "looks like a counterexample" to the theory of *-ing* clauses being the only type of complement allowing unspecified control. He explains that in the example sentence, the *to*-infinitive is different in function from the *to*-infinitive in the subject control construction "John tried to be careful." This shows that Duffley recognises two types of *to*-infinitives: those that allow unspecified control and those that only take subject control. He concludes that other verbs behaving similarly to *say* are also verbs of communication. He does not use *propose* as an example, but it does fall into the same category. As with *-ing* clauses, "the interpretation depend[s] on the lexical content of the matrix" (ibid: 242).

As regards the development of unspecified control over time, its use increases notably in the CLMET3: a total of 36 cases were identified as likely cases of unspecified control with the *to*-infinitive and two with the *-ing* clause (as opposed to four and two in the CLMET2). However, of these, 20 came from Booth's *In Darkest England and the Way Out*:

(31) I have already described how I *propose* to deal, in the first case, with the mass of surplus labour which will infallibly accumulate on our hands as soon as the Shelters are more extensively established and in good working order. But I fully recognise that when all has been done that can be done in the direction of disposing of the unhired men and women of the town, there will still remain many whom you can neither employ in the Household Salvage Brigade, nor for whom employers, be they registered never so carefully, can be found.
(CLMET3: 1890 Booth, *In Darkest England and the Way Out*)

Thus, in addition to the steep increase in *to*-infinitive complements, Booth's text plays a major role in the use of unspecified control in the CLMET3.

In the BNC unspecified control was identified in 12 *to*-infinitives and four *-ing* clauses:

(32) That was before Stonehenge, in 1980, became enclosed in a concentration camp cage, designed to keep people out, not in, and that fence was built before a Frenchman *proposed* to preserve the crumbling Sphinx by encasing him/her/it in a transparent plastic skin. (BNC: FET 2205)

(33) Well, someone got up and said I couldn't *propose* throwing redcoats out of Belfast because there weren't any nowadays. (BNC: A0U 419)

16 instances means that unspecified control occurred in 12% of all tokens of *to*-infinitive and *-ing* clause complements in the BNC. The number is 6% for the CLMET2 and 22% for the CLMET3. If Booth's text is left out of the calculations, the percentage is 15 for the CLMET3, which suggests that the major change happened between 1780 and 1920, as can be seen from the summary in Table 4. One reason for the increase of the unspecified control type might be the decrease of sentential complement patterns with the object of the proposal, i.e. *to* + NP, spelled out.

Table 4. Control in the *to*-infinitive and *-ing* clause complements of *propose* in the three subcorpora. The numbers exclude Booth's text

	Subject control		Unspecified control	
Subcorpus	**Raw count**	**Proportion**	**Raw count**	**Proportion**
CLMET2	92	94%	6	6%
CLMET3	100	85%	18	15%
BNC	123	88%	16	12%

If we return briefly to Duffley's (2000: 237) comment on the semantics of the matrix verb determining the understood subject, we can expect to find differences between the senses of *propose* presented in Section 2.1 in terms of control. In fact, in her research on "clauses [that] involve an unspecified or non-controlled interpretation of the missing subject", Fanego (2007: 178–186) looks at the history of *propose* and control. She found that in sense 1, "[t]o set before oneself as something to be done; to intend, purpose, or design",[6] *propose* used to take *to*-infinitive complements "and qualifie[d] as a subject-control verb" (ibid: 182). The *-ing*

6. These are the senses as given in the *OED Online*. Fanego's (2007) definitions are worded slightly differently, as she has used a different version of the *OED*.

clause complement began to be used once sense 2, "[t]o put forward or suggest as a scheme, plan, or course of action; to recommend or advocate that something be done", emerged in the 17th century. Fanego concludes that "[i]n the course of the Late Modern period and the twentieth century, the gerund steadily consolidated itself as a very frequent option for the coding of non-finite object complements without a specific controller" (ibid: 184).

Although, on the basis of the current study, Fanego's conclusion does not hold true for *propose* anymore, unspecified control does indeed occur almost exclusively with sense 2 in the three subcorpora:

(34) So I *propose* confining our attention to the elementary rules.
(CLMET3: 1909 Jerome, *They and I*)

Although control is closely related to the senses of *propose*, the increase of unspecified control in the current study cannot be pinpointed to the senses, as sense 2 has not increased in proportion to sense 1 between 1780 and the 1990s, as can be seen from Table 5. The increase of sense 2 in the CLMET3 is again explained by Booth's text.

Table 5. *To*-infinitive and *-ing* clause complements in relation to senses 1 and 2 in the subcorpora

	***to*-infinitive**		***-ing* clause**	
Subcorpus	**Sense 1**	**Sense 2**	**Sense 1**	**Sense 2**
CLMET2	52	22	17	5
CLMET3	91	61	6	6
BNC	103	23	7	3

5. Conclusions

The data suggest that major changes have taken place in the sentential complementation of *propose* between 1780 and 1920, after which the situation stabilised despite the verb becoming much less frequent overall. The *to*-infinitive has been the most common sentential complement of *propose* for the whole time period studied, and it strengthened its position further during this time. As the *-ing* clause complement has decreased in use, *propose* is an example of a predicate that goes against the general tendencies of the Great Complement Shift.

In addition to the *-ing* clause, all other sentential complement patterns have also decreased in use, not just in terms of NFs, but also in proportion to non-sentential complement patterns (see Figure 5). Exceptions to this are the

wh-clause, which only occurred once in both the CLMET3 and the BNC, and the quote, which did not occur in the CLMET3 at all, but reappeared in the BNC.

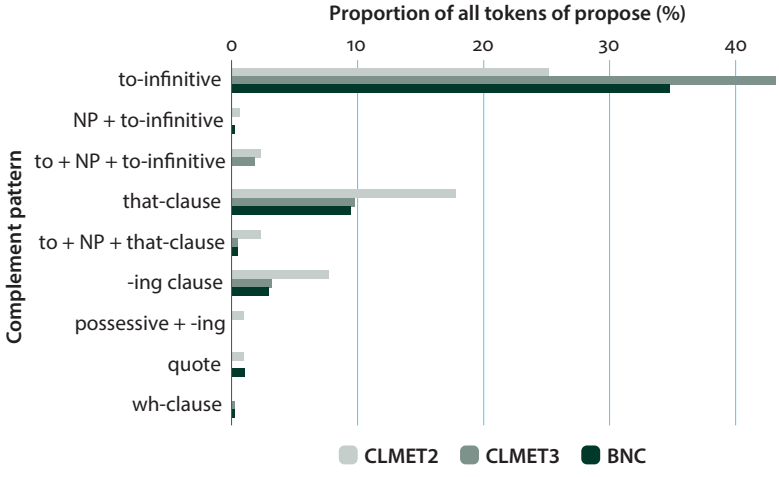

Figure 5. Changes in the proportions of the sentential complement patterns of *propose* in the subcorpora

The control structure of *propose* is diverse, as expected after the review of the literature in Section 2.3. Examples of all four control types were found in the data. Subject control is clearly the most frequent of the types, whereas object and PP object control are related to the infrequent complement patterns NP + *to*-infinitive and *to* + NP + *to*-infinitive. The latter pattern did not occur in the most recent data, which resulted in the loss of PP object control. Unspecified control increased in use over time following the pattern of the complements: a steep increase from the CLMET2 to the CLMET3 and stabilisation during the twentieth century.

The observations made here shed new light on the diachronic changes in the complement patterns as seen in British English. More work can still be done on the use of *propose*, using larger sets of data of different types of texts as well as regional varieties to establish a timeline for the diachronic development of the complement system of *propose* from the perspective of World Englishes – further research of this type is invited.

Acknowledgements

I would like to thank Juhani Rudanko for mentoring me in the writing of the MA thesis this article is based on. I am forever grateful for all the support and encouragement he has given me

throughout the years. I would also like to thank the two anonymous reviewers for their valuable feedback on the article.

References

Bach, Emmon. 1979. Control in Montague grammar. *Linguistic Inquiry* 19: 515–531.

Bach, Emmon. 1980. In defense of passive. *Linguistics and Philosophy* 3: 297–341. doi:10.1007/BF00401689

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

British National Corpus. <https://bncweb.lancs.ac.uk>

Corpus of Late Modern English Texts. <https://perswww.kuleuven.be/~u0044428/clmet.htm>

Davies, William D. & Dubinsky, Stanley. 2004. *The Grammar of Raising and Control: A Course in Syntactic Argumentation*. Oxford: Blackwell. doi:10.1002/9780470755693

Duffley, Patrick J. 2000. Gerund versus infinitive as complement of transitive verbs in English: The problems of "tense" and "control". *Journal of English Linguistics* 28(3): 221–248. doi:10.1177/00754240022005018

Fanego, Teresa. 2007. Drift and the development of sentential complements in British and American English from 1700 to the present day. In *'Of Varying Language and Opposing Creed': New Insights into Late Modern English* [Linguistic Insights Series 28], Javier Pérez-Guerra, Dolores González-Álvarez, Jorge L. Bueno-Alonso & Esperanza Rama-Martínez (eds), 161–235. Bern: Peter Lang.

Haegeman, Liliane. 1991. *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Huddleston, Rodney & Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Oxford English Dictionary Online. September 2013. Oxford: OUP. <http://www.oed.com>

Poutsma, H. MS. *Dictionary of Constructions of Verbs, Adjectives and Nouns*.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language* London: Longman.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17: 501–557.

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, style and grammatical norms*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166. Bern: Peter Lang.

Rudanko, Juhani. 1989. *Complementation and Case Grammar: A Syntactic and Semantic Study of Selected Patterns of Complementation in Present-Day English*. Albany NY: State University of New York Press.

Rudanko, Juhani. 2012. Exploring aspects of the Great Complement Shift, with evidence from the TIME Corpus and COCA. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Closs Traugott (eds), 222–232. Oxford: OUP.

Saarimäki, Veera. 2014. *The Complementation of the Verb Propose in Recent British English*. MA thesis. <https://tampub.uta.fi/handle/10024/96656>

Somers, Harold L. 1984. The validity of the complement-adjunct distinction in valency grammar. *Linguistics* 22: 507–530. doi:10.1515/ling.1984.22.4.507

Visser, F. Th. 1973. *An Historical Syntax of the English Language. Part Three, Second Half: Syntactical Units with Two and with More Verbs*. Leiden: E.J. Brill.

Vosberg, Uwe. 2003. The role of extractions and *horror aequi* in the evolution of *-ing* complements in Modern English. In *Determinants of Grammatical Variation in English* [Topics in English Linguistics 43], Günter Rohdenburg & Britta Mondorf (eds), 305–327. Berlin: Mouton de Gruyter. doi: 10.1515/9783110900019.305

Vosberg, Uwe. 2009. Non-finite complements. In *One Language Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 212–227. Cambridge: CUP. doi: 10.1017/CBO9780511551970.012

The use of optional complement markers in present-day English

The role of passivization and other complexity factors

Günter Rohdenburg
University of Paderborn

This paper reports on the results of a corpus-based study exploring several major types of omissible complement clause markers. Throughout, the distribution of rival variants with or without these markers is accounted for in terms of the Complexity Principle correlating processing complexity and grammatical explicitness. The focus is on the use of passivized matrix verbs as well as other complexity factors in the active, which favour the more explicit variants containing the function words in question. At the same time, the paper registers, wherever practical and instructive, any parallels and differences between British and American English. In most cases, the variation phenomena in question are the result of ongoing developments either introducing or reducing the use of the relevant function words.

Keywords: complement markers; variation; corpus linguistics; Complexity Principle; passivization; (in)transitive complement verbs

1. Introduction

The present paper focuses on the optional use of major dependent clause markers such as the complementizer *that*. In other words, as is seen with *tell* in (1a) and its passive counterpart (1b), we are dealing with grammatical functions that are either marked explicitly or left unmarked.

(1) a. Her friends told her (that) the proposal would not work.
b. She was told (by her friends) (that) the proposal would not work.

For various reasons, mainly to do with structural aspects and its reduced frequency of occurrence, the passive in English and other European languages has usually been regarded as a cognitively more complex category than the active (Wanner 2009), and its increased processing complexity has indeed been

DOI 10.1075/slcs.195.08roh

demonstrated or argued for by a succession of scholars (cf., e.g., Clark & Clark 1977: 105–106; Davison & Lutz 1985; Atkinson, Kilby & Roca 1988: 300–303, 309; Givón 1990: 957–958). Crucially, comparing actives and passives in cases like (1a)–(b), we find that the distribution of the two rival options differs between the active and the passive, with the (more complex) passive favouring the presence of the marker. This tendency is in line with the Complexity Principle, which has been explored in a series of studies (cf., e.g., Rohdenburg 1996, 2002, 2003, 2006a–b, 2007; Vosberg 2006; Mondorf 2009; Berlage 2014). The principle is usually stated as follows:

> In the case of more or less explicit constructional options, the more explicit ones(s) will tend to be preferred in cognitively more complex environments.

Additional support for this account is provided by the fact that – with a large variety of optional markers such as *that* – the more explicit alternative also tends to be promoted in the active by various complexities such as the following:

- structural discontinuities (e.g. Vosberg 2006)
- relatively complex clauses compared with the easy-to-process *there* clause (e.g. Rohdenburg 2003: 220–221, 2014b)
- noun phrase complexity including the pervasive contrast between full NPs and (personal) pronouns, in particular in subordinate clauses (e.g. Wasserman 1976; Berlage 2014; Rohdenburg 2014b)

The database used for this study consists of the LOB corpus and a large collection of full text British and American newspapers from the 1990s and early 2000s (cf. the bibliography for further details). Finally, concerning the corpus analyses conducted, a few methodological remarks may be in order:

- While the focus is on British English (BrE), comparisons with American English (AmE) or, exceptionally, explorations confined to AmE have usually been included where they have been possible and promised important insights.
- In order to guard against any potential distortions produced in the case of rival complement options by second and successive conjuncts, all analyses have been confined to the first in a series of conjuncts.
- Based on the chi square test, three levels of statistical significance will be distinguished: significant ($p < 0.05$ or $p < 0.025$), highly significant ($p < 0.01$) and extremely significant ($p < 0.001$). In the majority of sufficiently large samples, the figures compared reach significance levels ranging from $p < 0.01$ to $p < 0.001$. Therefore, explicit mention of the significance levels in question will usually be made only for tables containing what might appear to be doubtful cases.

– Both for reasons of economy and homogeneity, the search procedures adopted have been restricted in a wide variety of ways. In addition, in terms of size and text type, the datasets have generally been selected to suit the purpose in question as far as possible.

The rest of this article is organized as follows. Sections 2–6 survey four general kinds of clausal markers that are or have been optional with a range of governing expressions, the complementizer *that* after verbs and nouns (Sections 2.1 and 2.2), the infinitive marker *to* after verbs like *bid* and *assist* (Section 3), the modal *should* in complements after mandative predicates (Section 4), the item *from* introducing and marking gerundial complements after hitherto neglected transitive and intransitive particle verbs (Sections 5.1 and 5.2), and various other prepositions introducing interrogative complements (Section 6). No doubt, the presence of these clausal makers makes the relevant clausal structures more explicit, which largely explains their special affinity for the passive as well as various structural complications in the active. Finally, Section 7 provides a summary of the main findings of this paper and a general conclusion which stresses the continued relevance of work exploring the Complexity Principle's range of application.

2. The complementizer *that*

2.1 Verb-dependent complement clauses

With many verbs governing *that*-clauses, the complementizer can be omitted, and the choice between the more or less explicit grammatical variants is clearly influenced, among other things, by a range of complexity factors (cf. e.g. McDavid 1964; Wasserman 1976; Elsness 1984; Rohdenburg 1996; Tagliamonte & Smith 2005). This section uses two transitively used verbs, *tell* and *inform*, to investigate the extent to which the voice contrast and certain complexities in the active determine the distribution of the complementizer.

Let us begin by analyzing in Table 1 the verb *tell* in the LOB corpus. The analysis excludes any examples involving additional material following the object of *tell* and preceding the subject of the *that*-clause. The exclusion of these more complex examples in Table 1 is motivated by the fact that, in line with the Complexity Principle, the complementizer is here used almost categorically. In Table 1, a distinction is drawn – for the subject category of the *that*-clause – between personal pronouns and all other expressions in both the active and the passive.

Table 1. *that*-clauses associated with the verb *tell* in the LOB corpus

		I	II	III	IV
		that	**Ø**	**total**	**% *that***
1. active uses					
	a. all examples	93	84	177	52.5%
	b. *that*-clause subjects: personal pronouns vs all other items	53/40	71/13	124/53	42.7%/75.5%
	c. objects of *tell*: personal pronouns vs all other items	61/32	63/21	124/53	49.2%/60.4%
2. passive uses					
	a. all examples	18	6	24	75%
	b. *that*-clause subjects: personal pronouns vs all other items	4/14	6/0	10/14	40%/100%

The evidence in Table 1 provides a number of instructive findings. Most importantly, the proportion of the complementizer after the passive of *tell* is found to be significantly higher than that in the active (at $p < 0.05$). In addition, concerning the internal division of the subject category, we note a clear-cut contrast in both the active and the passive: personal pronoun subjects induce a strikingly lower incidence of *that* than all other subject expressions (at $p < 0.001$ in both the active and the passive). However, while personal pronouns in the object category also trigger the complementizer less frequently than all remaining NPs, the contrast fails to reach the basic significance level set at $p < 0.05$. Finally, the comparison of the two subject types in actives and passives suggests that the overall contrast between actives and passives is entirely due to the contrast involving NPs other than personal pronouns.

The second analysis is devoted to *inform* in the 1993 issues of the *Daily Mail*. *Inform* is a relatively formal verb whose complements could be expected to preserve a higher proportion of *that* than *tell.* The choice of the *Daily Mail* was motivated by the tendency of tabloids – observed, for instance, by Mondorf (2009) – to be more sensitive to complexity contrasts than quality papers, and the 1993 issue, by far the shortest of the available ones, was assumed to be sufficient for the purpose in hand. Again excluding any additional elements between the superordinate object and the subject of the *that*-clause, the data in Table 2 clearly corroborate the two prominent findings made in the study of *tell.* First, at $p < 0.05$, superordinate active clauses overall trigger a lower percentage of examples containing the complementizer than their passive counterparts. Second, at $p < 0.025$, personal pronoun subjects within the *that*-clause are less likely to retain the complementizer than the remaining and more complex NPs. However, due to the

overall much smaller contrast between complement clauses after *inform* with and without *that*, the two other comparisons – involving superordinate clause objects and subordinate clause subjects after passive *inform* – turn out to yield far from significant results.

Table 2. *that*-clauses associated with the verb *inform* in the 1993 issues of the *Daily Mail*

		I	II	III	IV
		that	**Ø**	**total**	**% *that***
1. active uses					
	a. all examples	128	17	145	88.3%
	b. personal pronoun subjects in the *that*-clause	58	13*	71	81.7%
	c. remaining NPs	70	4	74	94.6%
2. passive uses		55	1	56	98.2%

*The figure includes one example using *there* in the subject slot.

2.2 Noun-dependent complement clauses

Turning now to *that*-clauses governed by nouns, we will examine the contrast between *there*-clauses and all others, which, as a group, can be assumed to be more complex. With nouns governing interrogative clauses as in (2), it has been demonstrated that the choice of the more explicit prepositional option is influenced by a number of complexity constraints.

(2) There is uncertainty (about) whether the levy will apply.

For instance, in cases like (2) where the relevant noun is introduced by the common and easy-to-process *there*-clause, the prepositional option is selected less frequently than after corresponding nouns in the remaining clauses (Rohdenburg 2002: 86–87, 2003: 220–221). A further parallel is provided by *there*-clauses containing the antecedents of various kinds of relative clauses (cf. e.g. Tagliamonte & Smith 2005; Rohdenburg 2014b). Here, too, the zero variant is favoured in the environment of the *there*-clause.

Very similar contrasts can be observed with many nouns such as *risk* as in (3) which govern *that*-clauses offering a choice between the more explicit version including *that* and the less explicit one without it.

(3) There is also the risk (that) the United Kingdom will break apart.

Thus, the evidence in Table 3 shows that the less explicit option omitting *that* is favoured in the environment of easy-to-process *there*-clauses over the complement set of clauses, with the relevant figures for both BrE and AmE reaching

a significance level of $p < 0.01$. For quite some time, English has experienced a general trend towards dropping the complementizer *that* (cf., e.g., Rohdenburg & Schlüter 2009: 407–409). As is seen in Table 3, the development is usually more advanced in AmE than BrE. The data suggest that, in other than *there*-clauses, the erosion of the complementizer must be fairly recent. We may conclude, therefore, that the divergence between *there*-clauses and the remaining ones must have been evolving right from the start.

Table 3. *that*-clauses governed by and immediately following the singular noun *risk* in British and American newspapers (t97-t01, g97-g01, d97-d00, d02; L96–99)*

		I	II	III	IV
		that	Ø	total	% ***that***
1. British English					
	a. *there*-clauses	407	39	446	91.3%
	b. others	246	6	252	97.6%
2. American English					
	a. *there*-clauses	28	10	38	73.7%
	b. others	56	4	60	93.3%

*The analysis is confined to subject-initial *that*-clauses introduced by the definite article.

3. Infinitive marking

This section is concerned with the infinitive marker *to,* which in Middle English was used variably after a number of verbs. Crucially, even then, the more explicit marked infinitive tended to be less common after active verbs than passive ones (cf. e.g. Mustanoja 1960: 526, 529, 533). After a long period of vacillation, several verbs governing the sequence object + infinitive with or without the infinitive marker *to* have evolved a regular voice contrast in this respect (cf. the historical survey in Schlüter 2005: 185–206), with the active steadily reducing the proportion of the infinitive marker and the passive increasing it further. Thus causative *make* reached a near-stable state in the course of the 18th century by practically bringing a centuries-old evolution to a close. In present-day English, the verb exhibits a clear-cut active-passive asymmetry, with the marked and the unmarked (or bare) infinitive being either categorical or unacceptable after passive and active uses, respectively. By contrast, *help* is still on the move: while there is a definite trend in the active towards the zero variant in both AmE and BrE (see e.g. Mair 1995, 2002: 121–129; Rohdenburg 2009: 317–319; Berlage 2014), the passive may have been virtually confined to the marked infinitive for centuries. In the British

BNC, for instance, all of the relevant 167 passive examples of *help* contain the marked infinitive.

With both *help* and *make* in the active functioning as superordinate verbs, the earlier or present-day rivalry between the two infinitival variants has been shown to be motivated, among other things, by the Complexity Principle (for *make*, see Rohdenburg & Schlüter 2000: 445–451; Iyeri 2012: 65–67, and for *help*, see Rohdenburg 2000: 30–31; Berlage 2014). In addition, it is reasonable to assume that the establishment of the marked infinitive as the only option in the passive is largely due to its increased processing complexity (Schlüter 2005: 202).

We turn next to transitive uses of *bid* and *assist,* which have been observed to show a similar voice contrast. In the case of the directive verb *bid*, we are dealing with an archaic or obsolescent verb whose passive uses have also favoured the *to*-infinitive as in (4a)–(b) for the last few centuries (cf. e.g. Rohdenburg & Schlüter 2000: 485–488; Schlüter 2005: 203–206).

(4) a. I was bid to watch that you didn't go off out of the town, […] (NCF, 1861)
b. He was bidden to drive rapidly. (NCF, 1885)

By contrast, the active has generally tended to select the unmarked infinitive. Moreover, there is a further parallel between the general active-passive asymmetry and the contrast involving simple and more complex active uses: in line with the Complexity Principle, negated infinitives dependent on *bid* as in (5) have been found – in a 19th century narrative database – to systematically display a strikingly increased ratio of the infinitive marker over non-negated ones (Rohdenburg 2015: 111–112).

(5) When I bade you not to come to me in London because of my husband, I did not swerve. (NCF, 1869)

In addition, I carried out a further investigation of the verb *bid* in a large collection of British newspapers. This time, the analysis – which disregards one or two examples of *not*-negation, a phenomenon dealt with in Section 4 – has been confined to the verb forms *bade* and *bidden*. The exclusion of the remaining forms *bid, bids* and *bidding* was dictated by reasons of practicality: in present-day newspapers, these items account for thousands of examples and they are virtually always found in irrelevant senses or as homonymous instances of the noun *bid*. The evidence in Table 4 corroborates the clear-cut active-passive contrast already characteristic of earlier periods. Moreover, the distribution of the marked and the bare infinitive conforms to the expectations raised by the Complexity Principle. The object NPs in rows 1b-d have been assigned to three ascending degrees of (presumed) processing complexity, and the corresponding increases of

the marked infinitive from personal pronouns to other 1–2-word NPs and from the latter to 3–9-word NPs are found to be statistically significant at $p < 0.001$ and $p < 0.05$, respectively.

Table 4. The rivalry between marked and bare infinitives after *bid* (represented by active and passive *bade* and active and passive *bidden*) in British newspapers (t90-00, g90-00, d90-00, d02-04, m93-00, i93-94, i02-05)

		I	II	III	IV
		to-infinitive	bare infinitive	total	% ***to***
1. active uses					
	a. all examples	18	98	116	15.5%
	b. (unadorned) personal pronouns	3	75	78	4%
	c. remaining 1–2-word NPs	7	18	25	28%
	d. 3–9-word NPs	8	5	13	61.5%
2. passive uses		85	–	85	100%

Concerning the infinitive after the transitive use of *assist,* AmE, in particular, has in recent times begun to establish the zero variant as in (6) (Rohdenburg 2012: 149).

(6) Too many people throughout the region would like the satisfaction of being able to assist Detroit help itself. (D94)

Consider now the results of the extended analysis in Table 5, which is also restricted to two verb uses and for similar reasons to those that have been given in the case of *bid*.

Table 5. The rivalry between marked and bare infinitives after the verb *assist* (represented in the active by *assisting* and *to assist*) in American newspapers* (databases used: *assisting* + *to assist*: L97-00, W90-92, D92-95; passives: L92-99, D92-95, W90-92, N01)

		I	II	III	IV
		to-infinitive	bare infinitive	total	% ***to***
1. active uses					
	a. all examples	95	26	121	78.5%
	b. *assisting*	50	6	56	89.3%
	c. *to assist*	45	20	65	69.2%
2. passive uses		11	–	11	100%

*The passives searched for involve the perfect participle preceded by a standard non-contracted form of the verb *be* within a window of nine words.

Despite the rarity of relevant passive examples, we can probably assume that the unmarked infinitive has so far been confined to the active. Comparing the small passive total with the active total or either of the two active verb uses, we only find one significant contrast (at $p < 0.05$), that involving *to assist* as in (6). However, the greater advance of the bare infinitive associated with *to assist* may be largely due to a *horror aequi* effect, the tendency to avoid the coincidence of two (near-) adjacent marked infinitives (cf. e.g. Vosberg 2006; Rohdenburg 2003: 236–242; Berlage 2014).

We have seen, then, that while the evolutionary pathways of the four verbs under investigation may involve both minor and major differences, they agree in one important respect predicted by the Complexity Principle: wherever there is an evolving or stable active-passive asymmetry, it is the passive that displays a higher incidence of the marked infinitive.

4. Modal *should* in mandative complements

From an evolutionary perspective, modal *should*, which is part of the subordinate predicate in question, certainly represents an unusual complement marker. However, just like the infinitive marker, the complementizer *that*, and the grammaticalized use of *from* in Section 5, *should* is used optionally to explicitly identify functionally specific complement clauses. Similarly, the prepositions introducing interrogative complement clauses (as will be seen in Section 6) also function optionally to explicitly identify the clause as a complement.

Recent research has shown that for over a hundred years there has been a revival of the subjunctive in AmE, which has been spilling over into BrE (Kjellmer 2009; Hundt 2009; Schlüter 2009; Rohdenburg 2009). This section compares, in two case studies, the use of the subjunctive after mandative predicates as in (7a) with its counterpart in (7b) containing the additional modal *should*.

(7) a. She asked that the letter (not) be forwarded to Mr A.
b. She asked that the letter should (not) be forwarded to Mr A.

There are a number of reasons why the modal structure in (7b) should be regarded as representing a more explicit complement option than the subjunctive in (7a) (Rohdenburg 2015). These include the following two:

1. Unlike the subjunctive, which shares its form with the imperative, the infinitive and finite verbs, the *should*-structure unambiguously represents a finite form.
2. The subjunctive behaves syntactically like a reduced form of the *should*-structure in at least two respects: *not*-negation does not trigger *do*-support, and the negator occupies the position it has in the corresponding modal

alternative (cf. also Kjellmer 2009). Moreover, as pointed out by Dixon (1991: 38), "the meaning of the verb (sc. the superordinate mandative verb) implies 'should' in the complement clause".

The first case study examines the effect of *not*-negation on the distribution of the rival constructions. No doubt, the processing of the complement in (7a)–(b) is made more complex by the inclusion of *not*-negation (cf., e.g., Clark & Clark 1977: 107ff; Givón 1984: 321ff). Moreover, Horn (1978: 191–205) has shown that, cross-linguistically, negated complements tend to be realized by maximally explicit clausal structures. In line with Horn's Embedded Negation Constraint, which may be regarded as a special manifestation of the Complexity Principle, we would therefore expect the more explicit *should*-structure to be preserved much better in *not*-negated complements than elsewhere. The results in Table 6 confirm that this is indeed the case for BrE.

Table 6. Passive complements associated with and following the verb *ask* (used without a personal object) in British and American newspapers (t90-00, g90-00, i93-94, i02-04, m93-94; W90-92)*

		I	II	III	IV
		should	**subjunctive**	**total**	**% *should***
1. British English					
	a. *not*-negation	30	23	53	56.6%
	b. remaining cases excluding other forms of negation	60	435	496	12.3%
2. American English					
	a. *not*-negation	–	99	99	0%
	b. remaining cases excluding other forms of negation	–	176	176	0%

*The analysis is confined to sequences of the form *be* + V-*ed*/V-*n*.

In AmE, by contrast, the reintroduction of the subjunctive has – in this particular case – virtually phased out the modal alternative, thus leaving no scope for the Complexity Principle to apply. Significantly, the contrast between the negated use of modal *should* and the subjunctive parallels that between negated marked and bare infinitives after *bid* referred to in Section 3.

This brings us to the second case study involving the verb *demand* as in examples (8a)–(b).

(8) a. We are demanding that the shortfall be made good by MGN or any owner, […] (g91)

b. Britain countered by demanding that the precise terms of the deal should be negotiated after normalisation. (g91)

Informal observations had suggested to me that – in addition to *not*-negation – subject complexity in the complement clause might also play an important role in supporting the selection of the more explicit modal alternative. To simplify the search procedure, it was decided, therefore, to analyze just one verb form, namely *demanding* as in (8a)–(b), and to focus on subject expressions introduced by the definite article. The subsequent analysis summarized in Table 7 distinguishes between three degrees of complexity, minimal NPs containing two words (NP2), three-word phrases (NP3) and those featuring four and more words (NP4-NP9). The evidence conforms to the initial hypothesis: different degrees of subject complexity are seen to correlate with ascending proportions of the modal construction. More precisely, the following contrasts turn out to be statistically significant: NP2 vs NP4-NP9 at $p < 0.025$, NP2 vs NP3 + NP4-NP9 at $p < 0.05$, and NP2 + NP3 vs NP4-NP9 at $p < 0.05$. A corresponding analysis of American newspapers has not been undertaken for a very simple reason. Preliminary soundings suggest that the modal construction has been almost completely replaced by the subjunctive.

Table 7. The effect of subject complexity in the complement clause on the rivalry between modal *should* and the subjunctive after the verb form *demanding* in British newspapers (t90-00, g90-00, d91-00)*

	I	II	III	IV
	should	**subjunctive**	**total**	**% *should***
1. NP2	16	163	179	8.9%
2. NP3	12	82	94	12.8%
3. NP4-NP9	22	102	124	17.7.%

*The analysis is confined to passive complements of the form (*should*) *be* + past participle in subject-initial *that*-clauses.

5. The variable use of *from* introducing gerundial complements

5.1 Negative verbs of causation like *put off*

This section deals with several verbs of negative causation, which optionally use the item *from* to introduce a gerundial complement. Predictably, it is the more explicit structures including *from* that tend to display a higher frequency of occurrence in the passive than the active. Let us begin by sketching the relevant behaviour of

prevent, stop and *save*, which in recent times have evolved an intriguing voice split in BrE. Consider first examples like (9a)–(b) in present-day BrE, where the use of the preposition *from* would be obligatory across the board in AmE.

(9) a. These circumstances prevented/stopped/saved us (from) wasting too much money.
b. We were prevented/stopped/saved (by these circumstances) from wasting too much money.

For two reasons, the variants containing the preposition may be assumed to be more explicit than those without. First, the presence of *from* in examples like (9a)–(b) can be treated as an extension of the – once common – structure of preposition + NP to that of preposition + (gerundial) clause. Second, the use of *from* in (9a)–(b), which constitutes an immediately transparent transfer from the spatial or temporal domain to that of negation, explicitly marks the semantic orientation of the non-finite clause.

In contrast to AmE, BrE seems to have been reverting to an earlier state, which was compatible with both the prepositional and the zero option (Mair 2002: 112–115). Crucially, the revival of the less explicit zero variant in BrE has been confined to the active, which had been lagging behind the passive for centuries anyway (cf., e.g., Iyeri 2010: 87–99; Rohdenburg 2000: 36–37). To some extent, the choice between the two active options in (10a) – with or without the preposition *from* – is surely influenced by the complexity of the direct object (Rohdenburg 2000: 36–37 and Vosberg 2006: 149–157 on *prevent*, Rohdenburg 2002: 85–86 on *save*, and Rohdenburg 2006b: 63–64 on *stop*).

While the variable use of *from* – in BrE – with the gerunds of *prevent, stop, save*, and also *forbid* has at least been mentioned by a succession of grammarians, the largely similar distribution of *from* with the following particle verbs has generally been ignored up to now:

frighten off, put off, scare off, turn off, warn off

But then, large-scale corpus analyses (not detailed here) have shown that most of the relevant gerunds are relatively infrequent, in particular in AmE. Moreover, with two of these verbs, *frighten off* and *warn off*, the huge American newspaper database analyzed for this study does not yield a single gerund like those in (9). Owing to the general infrequency of relevant active and passive gerunds, only two verbs, *put off* and *warn off* (in BrE), lend themselves to a comparison of active and passive uses. With both verbs, the analysis distinguishes between the following three categories:

i. In the active, the choice of a comparatively complex object as in (10) triggers the ordering of *put* + *off* + object, thus separating the particle *off* from the

gerund. A similar separation of verb and gerund is brought about by the rare use of a *by*-phrase in the passive.

(10) These fears might put off many investors in rural areas ?(from) adopting such pension schemes.

ii. The canonical use in the active is represented by the ordering of *put* + object + *off* as in example (11), where the particle immediately precedes (the zero variant of) the gerund.

(11) We also find that the cost of buying the playing equipment puts people/them off (from) taking up the sport.

iii. In the passive, the canonical use without any intervening elements (with the latter including the *by*-phrase) has the particle immediately preceding the gerund as in example (12).

(12) I hope patients will not be put off (from) coming to our clinics because of this verdict.

Comparing the three ordering options in the newspaper data of Tables 8 and 9, we find that it is only structures like (10) separating the particle from the gerund where the additional preposition *from* occurs in all cases.

Table 8. The distribution of *from* with gerunds associated with the verb *put off* 'cause (s.o.) to dislike s.o. or something/discourage (s.o.) from doing something' in British and American newspapers (d91-00, m93-00; L92-99, D92-95, W90-02, N01)

		I	II	III	IV
		from	**Ø**	**total**	**% *from***
1. British English					
	a. *put* + *off* + object/*be* + *put* + *off* + *by*-phrase (1x)	28	–	28	100%
	b. *put* + object + *off*	4	≈ 416*	≈ 420	1.0%
	c. passive: *(be) put* + *off (from) -ing*	37	≈ 183*	≈ 220	16.8%
2. American English					
	a. *put* + *off* + object	–	–	–	–
	b. *put* + object + *off*	3	9	12	25%
	c. passive: (*be) put* + *off (from) -ing*	3	1	4	75%

*In this case, the grand total arrived at here has not been checked again. So chances are it may be out by one or two digits.

Table 9. The distribution of *from* with gerunds governed by the verb *warn off* in British newspapers (t90-01, g90-00, d91-00, m93-00)

	I	II	III	IV
	from	**Ø**	**total**	**% *from***
1. *warn* + *off* + object	11	–	11	100%
2. *warn* + object + *off*	1	45	46	2.2%
3. passive: (*be*) *warned* + *off* (*from*) *-ing**	10	37	47	21.3%

*This rubric contains three *-ing* forms which could be interpreted as lexicalized items, *racing* (1x) and *typing* (1x) in column I and *racing* (1x) in column II.

However, in the remaining 'canonical' actives and passives (i.e. examples [11] and [12]), where the particle *off* immediately precedes the gerund and is able to additionally function as a surrogate of *from,* we can observe a clear contrast between the active and passive uses in both BrE and AmE. *Put off* in the case of AmE and both *put off* and *warn off* in BrE exhibit the same kind of active-passive asymmetry in rows 1b–c and 2b–c of Table 8 and rows 2 and 3 of Table 9: while *from* is rarely found in the active, it occurs much more frequently in the passive. In Tables 8 and 9 the figures for BrE are extremely significant and highly significant at $p < 0.001$ and 0.01, respectively. However, those for AmE – apart from being insufficient – fail to reach the basic level of significance set at $p < 0.05$. Incidentally, the data in Table 8 suggest that here, too, AmE displays a stronger affinity for the additional use of *from* than BrE.

5.2 The particle verb *hold off*

We turn now to the class of intransitive verbs governing gerunds introduced by optional *from*. The class is illustrated by the particle verb *hold off* as in (13a)–(b), which may be characterized as representing an intransitive and less complex counterpart of verbs like *put off*.

(13) a. People will hold off from buying new cars towards the end of 1999, […] (g98)

b. The advice from the Consumers' Association is to hold off buying for the time being. (g98)

Comparing the two options in a large database of British and American newspapers, a previous study (Rohdenburg 2012: 151) shows that the regional contrast observed with transitive verbs of negative causation like *prevent* or *put off* is reversed in the case of intransitive *hold off*. While the more explicit variant featuring *from* occurs in 36.8% of all cases in BrE, the ratio in AmE amounts to only

7.5%. In other words, AmE appears to be more sensitive to a complexity contrast than BrE, a finding that is paralleled in other domains (e.g. Mondorf 2009; Berlage 2014). Moreover, the proportion of the *from*-variant is seen to be much smaller with *hold off* than with most transitive verbs of negative causation. In addition, we also find some superordinate transitive verbs that trigger higher rates of interrogative complement clause marking by means of prepositions than corresponding intransitive uses (Rohdenburg 2003: 217–219).

This raises the question of whether the contrast between transitive and intransitive verbs within complement clauses may also function as a constraint influencing the use of specific complement markers. In the case of *hold off*, this would mean that the *from*-variant should be attracted more strongly to transitive subordinate verbs than intransitive ones. The hypothesis has been tested in a large number of British newspapers from the late 1990s. Accordingly, the subordinate verbs dependent on *hold off* were classified into two categories, transitive uses containing a direct object as in (13a) and intransitive ones as in (13b). There are two cases where the subordinate verb governs a gerundial and an interrogative complement clause rather than a direct object. These examples have been excluded from consideration. At $p < 0.025$, the results in Table 10 clearly confirm the hypothesis. The transitive examples do indeed trigger a larger proportion of the more explicit *from*-variant than the remaining intransitive uses.

Table 10. The distribution of *from* with gerunds after the intransitive verb *hold off* in British newspapers (t96-00, g96-00, d96-00, m96-00)

	I	II	III	IV
	from	**Ø**	**total**	**% *from***
1. transitive uses	53	106	159	33.3%
2. intransitive uses	7	38	45	15.6%

6. Interrogative complement clauses

Over the last few centuries, most verbs, adjectives and nouns that govern prepositional phrases have extended the prepositional linkage to dependent interrogatives as well. This means that the clause is explicitly identified as a complement. However, in many cases such as (14), the preposition is still only optional, with various complications in the active determining the share of the more explicit option (Rohdenburg 2003: 205–249).

(14) This reminded her (of) what (job) she was meaning to do/how (miserable) she had felt at the time.

In addition to *there*-clauses diverging from the set of all other clauses and the transitive-intransitive contrast referred to in Sections (2.2) and (5.2), respectively, the constraints in the active that doubtless conform to the Complexity Principle include the following: number contrasts (ibid., pp. 223–225), discontinuous structures (ibid., pp. 225–228), and the contrast between more complex finite and less complex infinitival clauses (ibid., pp. 228–232). Furthermore, the interrogative items and functions have been shown to involve various hierarchies prompting different degrees of prepositional use (ibid., pp. 208–214). As far as the advance of the more explicit prepositional variants is concerned, *what*-clauses are clearly ranked above *how*-clauses. Furthermore, with both *what*- and *how*-clauses, combinations such as *what* + N and *how* as a premodifier of adjectives and adverbs intriguingly tend to lag behind the simpler uses of *what* and *how*. On the basis of the complexity effects listed above we would also expect passive interrogatives to trigger a higher incidence of the prepositional variant than active ones. The evidence summarized in Table 11 shows this indeed to be the case, and that for both *what*- and *how*-clauses.

Table 11. Dependent interrogative clauses associated with the transitive verb *remind* represented by *remind/reminds/reminded* in the active and introduced by *what* and *how* in the *Daily Mail* and the *Mail on Sunday* for 1993–2000*

		I	II	III	IV
		of/about	**Ø**	**total**	**% *of/about***
1. *what*					
	a. active	109	84	193	56.5%
	b. passive	25	2	27	92.6%
2. *how*					
	a. active	30	53	83	36.1%
	b. passive	16	11	27	59.3%

*The analyses are confined to finite interrogative clauses. In addition, they exclude the determiner use of *what* as well as the premodifier use of *how.*

7. Conclusion

The present paper surveys several kinds of optional complement markers in present-day English including the complementizer *that*, the infinitive marker, modal *should* in mandative complements, the item *from* introducing gerunds after

verbs like *put off* and *hold off* as well as various other prepositions linking interrogative complements to their superordinate verbs. Thus, we can in all cases distinguish between more or less explicit grammatical alternatives, those containing the marker and those dispensing with it. This paper assumes that the relevant contrasts may be largely accounted for in terms of the Complexity Principle, which posits a correlation between the degrees of processing complexity and grammatical explicitness concerned. Crucially, it can be shown that passives tend to increase grammatical explicitness in a number of ways in which they resemble active clauses involving various complications such as structural discontinuities, comparatively heavy subjects and objects, and transitive complement verbs rather than intransitive ones.

Admittedly, the characterization of the Complexity Principle (introduced in Section 1) overstates its range of application. While the principle accounts for literally dozens of grammatical variation phenomena in English and other languages, a few counterexamples have been discovered, where other, independently motivated, tendencies produce converse effects. For instance, such contrary effects have been observed with variably reflexive verbs like *disport* or *indulge* (Rohdenburg 2014a), where the inclusion of prepositional complements tends to promote the choice of the arguably less complex variant without the reflexive.

A further problem is associated with the – cognitively complex – extraction of postverbal elements out of complement clauses. Depending on the constructions involved, the processing load may be mitigated by two antagonistic and complementary strategies (Rohdenburg 2016): a) the use of a more explicit and typically more bulky variant conforming to the Complexity Principle and b) the use of a simpler and typically less explicit alternative, in line with the Domain Minimization Principle (Hawkins 1999, 2004). Moreover, there are also some specific variation phenomena that might be accounted for by both the Complexity Principle and some other constraint, which act in combination. A possible example is provided by transitive particle verbs, where the tendency for the particle to be placed adjacent to the verb in the case of complex objects has been explained as conforming to the Domain Minimization Principle (Lohse, Hawkins & Wasow 2004). Here, it would also be plausible to interpret the tendency in terms of the Principle of End-Weight (Eitelmann 2016) and the Complexity Principle. The case for the Complexity Principle could be argued as follows: The ordering variant involving the adjacency of the two constituents of the particle verb represents a more explicit grammatical structure than the alternative separating the two constituents. Crucially, we find that the more explicit variant is preferred over its alternative to the extent that the object is made more complex. And then there is also the case of the verb *spare*, whose striking complement behaviour has yet to be given an in-depth analysis.

Preliminary soundings suggest that, compared with verbs like *prevent* or *put off* (see Section 5.1) and *save* (Rohdenburg 2002: 85–86), *spare* displays two apparently antagonistic tendencies. In line with the Complexity Principle, full NP objects in the gerundial construction trigger a distinctly higher percentage of the (presumed) more explicit *from*-variant than personal pronoun objects. However, unlike *prevent, put off* and *save*, the *from*-variant with *spare* is much more common in the active than the passive.

In theory, it would have been possible to invoke a number of additional factors – system-internal and sociolinguistic/stylistic ones – which together with the Complexity Principle could be shown to constrain the use of the optional variants dealt with in this paper. A case in point is provided by interrogative complements (Section 6), where different *wh*-items trigger the use of optional prepositions to different degrees. However, the detailed treatment of such issues has been beyond the scope of this study.

The moral of all this is clear: in order to enhance our knowledge of how languages deal with cognitively complex structures, it is necessary to continue analyzing an increasing variety of – both structurally and lexically – novel data, which may allow us to delimit more sharply the range of application of the Complexity Principle and any competing explanatory principles involved. It is hoped that the present article has made a useful contribution to this goal.

References

Primary sources

d91-00, 02, 04-05 *The Daily Telegraph and Sunday Telegraph* on CD-ROM 1991–2000, 2002, 2004–2005. Chadwyck-Healey/ProQuest. (478,837,273 words)

D92-95 *Detroit Free Press* on CD-ROM 1992–1995. Knight Ridder Information Inc. (102,989,512 words)

g90-05 *The Guardian* (including *The Observer* 1994–2005) on CD-ROM 1990–2005. Chadwyck-Healey/ProQuest. (645,817,821 words)

i93-94, 02-05 *The Independent and The Independent on Sunday* on CD-ROM 1993–94, 2002–2005. ProQuest. (242,608,117 words)

L92-95 *The Los Angeles Times* on CD-ROM 1992–1995 Knight Ridder Information Inc. (320,016,164 words)

L96-99 *The Los Angeles Times* 1996–1999 (courtesy of The Los Angeles Times Editorial Library). (275,506,490 words)

LOB *The Lancaster-Oslo/Bergen Corpus* on CD-ROM 1961. ICAME. (1,000,000 words)

m93-00 *The Daily Mail and The Mail on Sunday* on CD-ROM 1993–2000. Chadwyck-Healey. (206,762,410 words)

N01 *The New York Times* on CD-ROM 2001. ProQuest. (52,132,979 words)

NCF *Nineteenth Century Fiction 1999–2000 Chadwyck-Healey*. (37,589,837 words)
t90-04 *The Times and The Sunday Times* on CD-ROM 1990–2004 Chadwyck-Healey/ProQuest. (729,848,339 words)
W90-92 *The Washington Times* (including *Insight on the News* 1990–1992) on CD-ROM 1990–1992. Wayzata Technology. (93,889,488 words)

Secondary sources

Atkinson, Martin, Kilby, David & Roca, Iggy. 1988. *Foundations of General Linguistics*, 2nd edn. London: Unwin Hyman.

Berlage, Eva. 2014. *Noun Phrase Complexity in English*. Cambridge: CUP. doi:10.1017/CBO9781139057684

Clark, Herbert & Clark, Eve. 1977. *Psychology and Linguistics: An Introduction to Psycholinguistics*. New York NY: Harcourt Brace Jovanovich.

Davison, Alice & Lutz, Richard. 1985. Measuring syntactic complexity relative to discourse context. In *Natural Language Parsing*, David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds), 26–66. Cambridge: CUP. doi:10.1017/CBO9780511597855.002

Dixon, Robert M. W. 1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: OUP.

Eitelmann, Matthias. 2016. Support for end-weight as a determinant of linguistic variation and change. *English Language and Linguistics* 20: 395–420.

Elsness, Johan. 1984. *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65: 519–533. doi:10.1080/00138388408598357

Givón, Talmy. 1984/1990. *Syntax: A Functional-Typological Introduction*, Vols. I–II. Amsterdam: John Benjamins. doi:10.1075/z.17

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75: 244–285. doi:10.2307/417261

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: OUP. doi:10.1093/acprof:oso/9780199252695.001.0001

Horn, Laurence R. 1978. Some aspects of negation. In *Universals of Language*, Vol. IV, Joseph Greenberg (ed.), 127–210. Stanford CA: Stanford University Press.

Hundt, Marianne. 2009. *Colonial lag, colonial innovation* or simply *language change?* In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 13–37. Cambridge: CUP. doi:10.1017/CBO9780511551970.002

Iyeri, Yoko. 2010. *Verbs of Implicit Negation and their Complements in the History of English*. Amsterdam: John Benjamins. doi:10.1075/z.155

Iyeri, Yoko. 2012. The complements of causative *make* in Late Middle English. In *Middle and Modern English Corpus Linguistics: A Multi-Dimensional Approach* [Studies in Corpus Linguistics 50], Manfred Markus, Yoko Iyeri, Reinhard Heuberger & Emil Chamson (eds), 59–74. Amsterdam: John Benjamins. doi:10.1075/scl.50.08iye

Kjellmer, Göran. 2009. The revived subjunctive. In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds.), 246–256. Cambridge: CUP. doi:10.1017/CBO9780511551970.014

Lohse, Barbara, Hawkins, John A. & Wasow, Thomas. 2004. Processing domains in English verb-particle constructions. *Language* 80: 238–261. doi:10.1353/lan.2004.0089

Mair, Christian. 1995. Changing patterns of complementation, and concomitant grammaticalisation, of the verb *help* in present-day English. In *The Verb in Contemporary English: Theory and Description*, Bas Aarts & Charles F. Meyer (eds), 258–272. Cambridge: CUP.

Mair, Christian. 2002. Three changing patterns of verb complementation in Late Modern English: A real-time study based on matching text corpora. *English Language and Linguistics* 6: 105–131. doi:10.1017/S1360674302001065

McDavid, Virginia. 1964. The alternation of "that" and zero in noun clauses. *American Speech* 39: 102–113. doi:10.2307/453112

Mondorf, Britta. 2009. *More support for more-support: The Role of Processing Constraints on the Choice between Synthetic and Analytic Comparative Forms* [Studies in Language Variation 4]. Amsterdam: John Benjamins. doi:10.1075/silv.4

Mustanoja, Tauno F. 1960. *A Middle English Syntax, Part I: Parts of Speech*. Helsinki: Societé Néophilologique.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7: 149–182. doi:10.1515/cogl.1996.7.2.149

Rohdenburg, Günter. 2000. The complexity principle as a factor determining grammatical variation and change in English. In *Language Use Language Acquisition and Language History: (Mostly) Empirical Studies in Honour of Rüdiger Zimmermann*, Ingo Plag & Klaus Peter Schneider (eds), 25–44. Trier: Wissenschaftlicher Verlag.

Rohdenburg, Günter. 2002. Processing complexity and the variable use of prepositions in English. In *Perspectives on Prepositions*, Hubert Cuyckens & Günter Radden (eds), 79–100. Tübingen: Niemeyer. doi:10.1515/9783110924787.79

Rohdenburg, Günter. 2003. Cognitive complexity and *horror aequi* as factors determining grammatical variation and change in English. In *Determinants of Grammatical Variation in English*, Günter Rohdenburg & Britta Mondorf (eds), 205–249. Berlin: Mouton de Gruyter. doi:10.1515/9783110900019.205

Rohdenburg, Günter. 2006a. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms: English from 1500–2000*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166. Bern: Peter Lang.

Rohdenburg, Günter. 2006b. Processing complexity and competing sentential variants in present-day English. In *Linguistik International: Festschrift für Heinrich Weber*, Wilfried Kürschner & Reinhard Rapp (eds), 51–67. Lengerich: Pabst Science Publishers.

Rohdenburg, Günter. 2007. Functional constraints in syntactic change: The rise and fall of prepositional constructions in Early and Late Modern English. *English Studies* 88: 217–233. doi:10.1080/00138380601042824

Rohdenburg, Günter. 2009. Grammatical divergence between British and American English in the nineteenth and early twentieth centuries. In *Current Issues in Late Modern English*, Ingrid Tieken-Boon van Ostade & Wim van der Wurff (eds), 301–329. Bern: Peter Lang.

Rohdenburg, Günter. 2012. Britisches und Amerikanisches Englisch: Eine Sprache, zwei Grammatiken? In *Sprachmythen – Fiktion oder Wirklichkeit?*, Lieselotte Anderwald (ed), 137–160. Frankfurt: Peter Lang.

Rohdenburg, Günter. 2014a. Reflections on reflexives in Modern English. *Anglia* 132: 536–574. doi:10.1515/ang-2014-0057

Rohdenburg, Günter. 2014b. Relative clauses of reason in British and American English. *American Speech* 89: 288–311. doi:10.1215/00031283-2848978

Rohdenburg, Günter. 2015. The embedded negation constraint and the choice between more or less explicit clausal structures in English. In *Perspectives on Complementation: Theory, Structure and Variation*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 101–127. Houndmills: Palgrave Macmillan.

Rohdenburg, Günter. 2016. Testing two processing principles with respect to the extraction of elements out of complement clauses in English. *English Language and Linguistics* 20: 463–486.

Rohdenburg, Günter & Schlüter, Julia. 2000. Determinanten grammatischer Variation im Früh- und Spätneuenglischen. *Sprachwissenschaft* 25: 443–496.

Rohdenburg, Günter & Schlüter, Julia. 2009. New departures. In *One Language Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 364–423. Cambridge: CUP. doi: 10.1017/CBO9780511551970.020

Schlüter, Julia. 2005. *Rhythmic Grammar: The Influence of Rhythm on Grammatical Variation and Change in English*. Berlin: Mouton de Gruyter. doi: 10.1515/9783110219265

Schlüter, Julia. 2009. The conditional subjunctive. In *One Language Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 277–305. Cambridge: CUP. doi: 10.1017/CBO9780511551970.016

Tagliamonte, Sali A. & Smith, Jennifer. 2005. *No momentary fancy*! The *zero* 'complementizer' in English dialects. *English Language and Linguistics* 9: 289–309. doi: 10.1017/S1360674305001644

Vosberg, Uwe. 2006. *Die große Komplementverschiebung. Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.

Wanner, Anja. 2009. *Deconstructing the English Passive*. Berlin: Mouton de Gruyter. doi: 10.1515/9783110199215

Wasserman, Robert David. 1976. *Theories of Linguistic Variation*. PhD dissertation, Indiana University.

Patterns of direct transitivization and differences between British and American English

Marcus Callies
Universität Bremen

Rohdenburg (2009) found that prepositions are increasingly omitted in several types of verbs, marking a shift in complementation from an intransitive pattern with a prepositional object to a transitive pattern featuring a direct object noun phrase. In particular, the decrease of prepositional objects after antagonistic verbs (*appeal, battle, fight, protest*) and verbs of leaving (*depart, escape, flee, resign*) has been interpreted in line with an ongoing tendency in the history of English to functionally expand the category of the direct object at the expense of prepositional phrases in particular. Other verbs that are said to allow the preposition-less variant are mentioned only sporadically in the literature but have not yet been examined systematically on a broader empirical basis. This chapter provides corpus data from British and American English suggesting that several other verbs may also be in the process of undergoing direct transitivization.

Keywords: verb complementation; direct transitivization; omission of prepositions; grammatical economy; British English; American English

1. Introduction: Direct transitivization

The use of prepositions in various grammatical patterns has undergone significant changes in the (more recent) history of English. For instance, it has been observed that in a number of verbs, prepositions are increasingly omitted, resulting in a shift in complementation from an intransitive pattern with a prepositional object (PO) to a transitive pattern featuring a direct object (DO) noun phrase. This variation is exemplified in (1)–(5). The examples are taken from the *Corpus of Contemporary American English* (COCA; Davies 2008-) with optional prepositions representing the PO-variant added and given in brackets.

(1) There is plenty viewers can do to **protest** [against] this trend.
(COCA, 1998, *USA Today*)

DOI 10.1075/slcs.195.09cal

(2) While most people **flee** [from] these extreme weather conditions, there are those reporters and cameramen who boldly venture straight into the eye of the storm. (COCA, 2004, *CNN*)

(3) He **graduated** [from] Harvard summa cum laude. (COCA, 1993, *NPR Weekend*)

(4) For first-timers, **shopping** [at] IKEA can be daunting. (COCA, 1990, *Money*)

(5) Government regulation is another critical area **impacting** [on/upon] competitiveness. (COCA, 1990, *Money*)

This phenomenon has also been referred to in the research literature as "direct transitivization". Apparently it was Kirchner (1955) who coined the term (he referred to it as "direct transitivation"), although the alternation between a PO and DO had been noted much earlier, e.g. by Jespersen (1927; referred to in Kirchner 1955: 17) and Galinsky (1952: 161ff.). Early treatments mostly consist of brief discussions of assumed differences between British English (BrE) and American English (AmE), sometimes supplemented by extensive lists of verbs supposed to allow the alternation (see esp. Kirchner 1959, 1972 for such lists). However, there exist only very few empirical, corpus-based analyses of individual verbs (e.g. Hundt 1998 on *appeal* and *protest*). To my knowledge, the only substantial study focusing on a set of verbs and grammatical differences between BrE and AmE in NP-complements is Rohdenburg (2009). He situates his study in the context of the long-term consequences of the erosion of the English case system and the "freezing" of word order which brought about a tendency to expand the category of the direct object at the expense of the prepositional object. As a consequence, the category of the direct object became much more extensive and abstract. Rohdenburg analysed two semantic classes of verbs in detail: antagonistic verbs (*fight, battle, protest, appeal, race, play, offend*) and verbs of leaving (*flee, resign, escape, depart*). In a corpus study based on large corpora of British and US-American newspapers he found that in the use of the verbs under study POs are replaced by DOs to varying degrees (he observed "distinct distributional profiles") and that AmE is always further advanced than BrE, the contrast being especially sharp with *appeal* and *protest*. Rohdenburg thus concluded that with the said classes of verbs the DO has become the preferred variant in AmE in the course of the 20th century, while BrE still uses the PO variant almost exclusively.

Other verbs that are said to allow the alternation between the PO and DO variant are identified, mentioned and discussed very briefly in the more recent literature (see e.g. Algeo 1988, 2006), but have not yet been examined systematically on a broader empirical basis. For example, Algeo (2006) notes a tendency for AmE

to omit prepositions in *vote (for the) Democratic (Party), shop (at) Macys, fly (by) United, joke (with) someone.* Similarly, Rohdenburg (2009: 198, footnote 5) lists other verbs that corpus analyses supposedly have shown to involve similar contrasts between the two varieties of English: *cater (for/at), hunt (in/on etc.), impact ([up]on), shop (at/in), trail (behind/against), trample ([up]on/over)* and *work (at/in/on) a job*[1]. But to the best of my knowledge, there are no detailed empirical analyses yet.

To complicate this sketchy information, there are also different views as to the diachronic development of this variation and whether it is more prominent in BrE or AmE. Algeo (2006: 217) maintains that "a number of verbs in contemporary British take a nominal complement, whereas in American (and older British) use, they would normally have a prepositional complement instead". As mentioned above, he then lists a number of verbs that "can also be used with a noun phrase direct object in British" while at the same time arguing that "in American too there has been a recent tendency to omit prepositions" (2006: 218). More generally, Algeo suggests that this "innovative pattern is shared by the two national varieties, but the specific realizations of the pattern often differ" (2006: 218). Rohdenburg (2009) also seems to be arguing in favour of an innovative pattern in AmE by showing that the DO variant is more typical of AmE with the set of verbs he examined. He suggests that AmE is "clearly further advanced than BrE".

In earlier research, Galinsky (1952) and Kirchner (1955, 1959) had argued that direct transitivization was a much older phenomenon. Kirchner (1955) proposed that the preposition-less variant could in fact be dated back to Early Modern English and that it had been retained to a much greater extent in general AmE usage, while except for several verbs of movement, remnants of this variant in BrE could only be found in technical jargon or slang (Kirchner 1955: 18, footnote 3). Kirchner's argument in favour of "an American predilection" for direct transitivization (1955: 22) could thus be interpreted as a case of 'colonial lag', i.e. an apparent retention of older mother-country (i.e. British) usage in present-day AmE. Similarly, Galinsky (1952) argued that syntactically, AmE appears to be more conservative in the use of prepositions, retaining many older transitive uses (1952: 164f.; 192ff.).

This chapter aims to add to the empirical evidence on direct transitivization by providing the results of a case study that focuses on three verbs that have been claimed to be subject to direct transitivization: *graduate*, *impact*, and *shop*. These

1. Unfortunately, Rohdenburg does not provide references to point the reader towards the said analyses.

verbs were chosen because they are claimed most frequently in the pertinent literature to allow the alternation between the PO and DO variant. However, these claims are typically based on limited, often anecdotal evidence and thus, they need to be checked against actual usage data by means of large corpora. On the basis of evidence from dictionaries and corpora of British and American English, the following research questions will be examined:

1. Is there variation between the PO and the DO variant, and if so, how substantial is it in the three verbs under study? Which variety shows more variation, BrE or AmE?
2. Are there differences as to the preferred variant between AmE and BrE?
3. What is the evidence for the diachronic development of this variation in the 20th and early 21st centuries?

2. Lexicographic treatment of the different complementation patterns

Although Kirchner already noted in the 1950s that the phenomenon under study had been largely overlooked by grammarians (1955: 17), it still comes as a surprise that the variation between the PO and DO variants is apparently not covered in modern, corpus-based reference grammars of English. Herbst et al.'s (2004) valency dictionary of English lists only few of the verbs examined by Rohdenburg (2009) that allow the alternation, e.g. *escape*, *fight* and *protest*. To check if the transitive use with the DO variant has been documented by lexicographers, three major dictionaries of English were consulted, i.e. the *Longman Dictionary of Contemporary English* (LDOCE), the *Oxford English Dictionary* (OED), and for American English usage *Merriam-Webster's Learner's Dictionary* (MW). All dictionaries were used in their latest and presumably most up-to-date online versions. Having said that, the entries for *graduate* and *impact* provided by the OED may no longer reflect contemporary use because they date back to the late 19th and early 20th century respectively. However, the quotation database includes examples that are more recent, thus it remains unclear when exactly these OED entries were last updated. The entry for *shop* was updated only recently, i.e. in September 2016.

2.1 *Graduate*

In Table 1, the information for the verb *graduate* with the meaning 'obtain a degree from college or university' is compiled on the basis of the three major dictionaries mentioned above.

Table 1. *Graduate, v.* in three dictionaries of English

	pattern	meaning	example (if given)
LDOCE	intransitive + *from*, *in*	'obtain a degree, especially a first degree, from a college or university'	*Kate graduated from medical school last year. / He graduated in physics from Cambridge University.*
	intransitive + *from* (AmE)	'complete your education at high school'	*Jerry graduated from high school last year.*
	transitive (esp. AmE)	'give a degree or diploma to someone who has completed a course'	
OED	intransitive	'take a university degree. Also (U.S.): 'complete a high school course and receive a diploma'	*In 1837 he graduated from Yale College.*
	transitive (now rare except U.S.)	'admit to a university degree. Also with complement, indicating the degree obtained'	*The class of '76 was graduated with six men.*
MW	intransitive*	'earn a degree or diploma from a school, college, or university'	*He graduated from the university last June. /They both graduated with honors.*
	transitive (US, of a school, college, or university)	'award a degree or diploma to (a student) - usually used as (*be*) *graduated*'	*He was graduated from the university last June.*
	transitive (US, informal)	'earn a degree or diploma from (a school, college, or university)'	*He joined the navy after graduating high school.*

*For this pattern, *Merriam Webster's Learner's Dictionary* adds that in BrE, *graduate* refers only to earning a college or university degree, while in AmE it is also used for other schools (such as high schools).

All three dictionaries list at least one transitive pattern that is explicitly marked as AmE usage. Overall, three transitive patterns can be identified, exemplified below by means of data taken from the COCA and the *British National Corpus* (BYU-BNC, Davies 2004-). Optional prepositions representing the PO-variant are given in brackets.

(6) He **graduated** [from, at] **Harvard** summa cum laude.
(COCA, 1993, *NPR Weekend*)

(7) He **graduated** [with, as] **MA** at Queen's college, Belfast and entered the Ministry in 1874. (BNC, B13 486)

(8) a. Stanford University yesterday **graduated 90 new doctors** …
(COCA, 1992, *San Francisco Chronicle*)

b. **He was graduated** last June with a degree in economics …
(COCA, 1992, CBS *Sunday Morning*)

2.2 *Impact*

Table 2 lists the lexicographic information for the verb *impact* 'have a strong effect on'. While both the LDOCE and the MW list the transitive DO variant without *on*/*upon*, this is absent from the OED.

Table 2. *Impact, v.* in three dictionaries of English

	pattern	**meaning**	**example (if given)**
LDOCE	intransitive and transitive (esp. AmE) + *on*, *upon*	'have an important or noticeable effect on someone or something'	*The Food Safety Act will progressively impact on the way food businesses operate.*
OED	intransitive	'have a (pronounced) effect on'	*The Magazine… is not the place for consideration of national and international events except in so far as they impact on Oxford.*
MW	intransitive	'have a strong and often bad effect on (something or someone); affect'	*The poor economy is impacting on/upon small businesses.*
	transitive	same meaning	*No one is sure how these changes will impact our relations with other countries.*

2.3 *Shop*

Table 3 summarizes the dictionaries' entries for the verb *shop* 'go to one or more shops to buy things'. Both the OED and the MW record a transitive use of *shop* which is labelled as AmE use and the meaning of which appears to be slightly different from the intransitive use of the verb. Excluded from this overview are two variety-specific transitive uses of *shop*: 1) 'try to get a company to publish or produce (something)' as in *She's shopping her idea for a film*, often used with the preposition *around* (*He shopped his manuscript around, but no publishers were interested*) and marked as AmE usage in both the OED and the MW; and 2) 'tell the police about someone who has done something illegal' as in *He was shopped by his ex-wife*, marked as informal BrE use in both the LDOCE and MW.

It can be summarized that in contrast to the major reference grammars of Present-Day English, the variation to be studied in this chapter has not escaped the attention of lexicographers as both the intransitive PO variant and the transitive DO variant are listed for all three verbs in at least one major dictionary. Moreover, and in accordance with the research literature, for two of the three verbs the transitive variant is marked as typical of AmE usage. The exact nature and extent of this variation will now be examined by means of data drawn from large electronic corpora of English.

Table 3. *Shop, v.* in three dictionaries of English

	pattern	meaning	example (if given)
LDOCE	intransitive + *for*, *at*	'go to one or more shops to buy things'	*I usually shop for vegetables in the market. / She always shops at Tesco's.*
OED	intransitive	'visit a shop or shops to buy or view goods; to examine or search for goods or services with intent to buy'	*The National Portrait Gallery went shopping at Phillips sale room yesterday.*
	transitive (orig. North American)	'go shopping at (a store, an online retailer, etc.); to examine goods on sale in (a shop, etc.)'	*From this month, Brits can shop the online store.*
MW	intransitive	'visit places where goods are sold in order to look at and buy things'	*I like to shop at locally owned stores.*
	transitive (US)	'examine the stock or offerings of'	*They shopped the store(s) in search of gift ideas.*

3. Data and methodology

While the so-called Brown family of corpora (see Xiao 2008: 395ff.) has been a major empirical resource for researchers interested in differences and changes in British and American English, each one of these corpora comprises only one million words of written language and is thus much too small for the present study. Instead, two large reference corpora were examined: 1) the *British National Corpus* (BNC), consisting of 100 million words of written (90%) and spoken (10%) language that represent British English of the early 1990s (see CoRD 2016), and 2) the *Corpus of Contemporary American English* (COCA), a monitor corpus covering American English from 1990 to 2015, currently including 520 million words. Because these two corpora do not exactly match for time period, it was decided to use only parts of the COCA as a basis for comparison to the BNC. The large majority of texts in the BNC (more than 90%, see Burnard 2007, Section 1.3) were published between 1985 and 1993, and thus, the search in the COCA was restricted to the earliest time span available, i.e. 1990–1994 (approx. 104 million words) so as to not include much data published after the compilation of the BNC.

In addition, to find further evidence of more recent, potential innovative uses of the verbs under study, and to examine the extent of the variation on the basis of more recent AmE data, I also searched the 2015 section of the COCA (approx. 20 million words). Another resource that provides more recent data for both BrE and AmE is the *Global Web-based English Corpus* (GloWbE; Davies 2013-,

Davies & Fuchs 2015), a web-derived corpus composed of 1.9 billion words from 1.8 million web pages in twenty different English-speaking countries (native English and English as a Second Language). The texts for the corpus were collected in December 2012 and mostly consist of informal blogs (about 60%) and other written texts harvested from the Internet, such as newspapers, magazines, and company websites (Davies & Fuchs 2015: 3). The fact that the corpus is web-derived makes it necessary to ensure that the webpages are correctly associated with each of the twenty countries represented in the corpus. To do so, the compilers collected the texts for each country separately, using the Google "Advanced Search" facility and limiting the searches by region. While it is not entirely clear how well Google has correctly identified websites by country (Davies & Fuchs 2015: 4) there are undoubtedly cases where language data appear on a webpage sorted into a particular country category although the data were not produced by a writer/speaker who speaks the particular variety associated with that country. Researchers are provided with the URLs for each of the 1.8 million webpages with hyperlinks to the original sources provided by the corpus interface, but background checks are sometimes difficult because the source websites have disappeared, or because manual examination of all the original webpages to identify the country of origin of the author/s cannot be carried out for high-frequency phenomena. Keeping these limitations in mind, the GB and US components of the GloWbE were searched to update and complement the findings obtained from the BNC and COCA data.

Moreover, to find evidence for the development of the variation in a wider diachronic perspective, the *Corpus of Historical American English* (COHA; Davies 2010-) was used. It covers written AmE from 1810 to 2009. For the present study, the search was restricted to the time span of 1900–2009.

All pertinent forms of the respective verb lemmas were extracted from the corpora and then manually inspected with false positives being discarded. Tokens with intervening material between a PO or DO and the verb (e.g. *graduated* ***instead*** *from*, *impact* ***most heavily*** *on*) were analysed separately.

4. Results

4.1 *Graduate*

For this verb, three transitive patterns mentioned earlier were examined (again, optional prepositions representing the PO variant are given in brackets):

– **Pattern 1:** He **graduated** [from, at] **Harvard** summa cum laude.
(COCA, 1993, *NPR Weekend*)
– **Pattern 2:** He **graduated** [with, as] **MA** at Queen's college, Belfast and entered the Ministry in 1874. (BNC, B13 486)

– **Pattern 3a (active):** Stanford University yesterday **graduated 90 new doctors** … (COCA, 1992, *San Francisco Chronicle*)
– **Pattern 3b (passive): He was graduated last June** with a degree in economics … (COCA, 1992, *CBS Sunday Morning*)

The corpus data show that Pattern 1 is clearly the most frequent one, but that the DO variant is not attested in late 20th century BrE and is only a minority variant in AmE, see Table 4.

Table 4. Pattern 1: *graduate from university / at Oxford* vs. *graduate college / Yale*.

	N	PO *(from, at)*	%	DO (∅)	%	+ ADV	% PO
BNC	140	140 (1.44 pmw[2])	**100**	0 (0 pmw)	0	29	**100**
COCA 1990–1994	1157	1113 (10.7 pmw)	**96.2**	44 (0.4 pmw)	3.8	62	**100**

The total count of all instances of the PO variant (140 in the BNC, 1113 in the COCA) includes those in which an adverbial follows the verb (29 in the BNC and 62 in the COCA), see examples (9) and (10). It is important to note that the occurrence of an intervening adverbial invariably triggers the PO variant.

(9) Cecil **graduated last year from the University of Michigan**. (COCA, 1992, *ABC Special*)

(10) He dropped out of Southeast Missouri State and **graduated instead from the Elkins Institute of Radio and Technology in Dallas**. (COCA, 1994, *Washington Post*)

With Pattern 2, perhaps surprisingly, it is BrE that favours the DO variant, while AmE overwhelmingly uses the PO variant, see Table 5.

Table 5. Pattern 2: *graduate as/with (a) BA* vs. *graduate MA*

	N	PO *(with, as)*	%	DO (∅)	%
BNC	40	8 (0.08 pmw)	20	32 (0.3 pmw)	**80**
COCA 1990–1994	14	13 (0.1 pmw)	**92.9**	1 (0.001 pmw	7.1

Finally, Pattern 3 is the least frequent one in general. It is practically absent from BrE and shows a clear preference for the active variant in AmE (see Table 6).

2. pmw = normalised frequency per million words.

Table 6. Pattern 3: *graduate 2,000 students* vs. BE *graduated in June / from college*

	N	active	%	passive	%
BNC	1	1 (0.01 pmw)	**100**	0 (0 pmw)	0
COCA 1990–1994	61	53 (0.51 pmw)	**86.9**	8 (0.08 pmw)	13.1

The corpus data thus confirm the information found in the dictionaries (see Table 1 above) in that two of the three transitive uses of *graduate* (Patterns 1 and 3) are specific to AmE. This is also confirmed by an analysis of all three patterns in the most recent section of the COCA (2015). Table 7 illustrates that while the PO variant is clearly the predominant variant for Patterns 1 and 2, we do find a fair degree of variation in the first pattern (*graduate high school / college / university*; 14.1%) and also the transitive use in Pattern 3 (*graduate someone*).

Table 7. All three patterns of *graduate* in the 2015 section of the COCA

	N		%		%	+ ADV	% PO
Pattern 1	375	PO (*from, at*) 322 (16.14 pmw)	**85.9**	DO 53 (2.66 pmw)	14.1	26	**100**
Pattern 2	53	PO (*with, as*) 52 (2.60 pmw)	**98.1**	DO 1 (0.05 pmw)	1.9	5	**100**
Pattern 3	15	active 15 (0.75 pmw)	**100**	passive 0 (0 pmw)	0	0	**0**

Again, the count for all instances of the PO variant includes cases with an intervening adverbial that invariably triggers this pattern.

An examination of the most frequent pattern *graduate college* in the COHA (see Table 8) suggests that we are dealing here with a fairly recent type of usage that only seems to have emerged in the course of the 20th century in AmE.

Table 8. *Graduate* used with the DO variant in pattern 1 (only instances where the verb is followed by a proper name or *college*, *university*, *high school*)

corpus	corpus size	DO	DO pmw
COHA 1900s–1940s	119,386,620	0	0.00
COHA 1950s	24,398,180	2	0.08
COHA 1960s	23,927,982	2	0.08
COHA 1970s	23,769,305	2	0.08
COHA 1980s	25,178,952	3	0.12
COHA 1990s	27,877,340	17	0.61
COHA 2000s	29,479,451	22	0.75

Whereas this pattern is not attested in the BNC, the more recent data harvested from the web suggest that in the course of the early 21st century the AmE transitive use is slowly making its way into BrE usage.[3] Because of the sheer size of the GloWbE it was impossible to analyse all forms of *graduate* in context exhaustively. Thus, I retrieved all occurrences of the verb when it was either directly followed by the preposition *from* or one of the most frequent direct object NPs *college*, *university*, *high school* and a proper name like *Stanford* or *Harvard* when used in the DO variant. The results are shown in Table 9.

Table 9. *Graduate* used with *from* vs. the DO variant in Pattern 1 in the GB and US sections of the GloWbE (only instances where the verb is followed by *college*, *university*, *high school* or a proper name)

corpus	corpus size	GRADUATE	+ *from*	+ DO	DOpmw
GloWbE GB	387,615,074	6832	1757	*high school* 41 *college* 30 *university* 6 proper name 13 total 90	0.23
GloWbE US	386,809,355	12525	4123	*high school* 412 *college* 246 *university* 15 proper name 53 total 726	1.88

There is a lexical bias towards a higher occurrence of the DO variant in AmE in that *high school* and *college* are generally used more frequently in AmE because of differences in the educational systems of the US and the UK (e.g., *high school* is typically used more generally for secondary school, and *college* is often used as a synonym for *university* in the US). Nevertheless, the data show that the DO variant also occurs with these nouns in BrE usage.

4.2 *Impact*

For the verb *impact*, the comparison of data from the BNC and the COCA suggest that the transitive use is more typically used in AmE as it is the preferred variant, whereas as the PO variant is still more frequent than the DO variant in BrE (see Table 10).

3. Note that due to the limitations of the web-derived data in the GloWbE, some of these occurrences that appear on webpages in Great Britain may have been produced by speakers of AmE (e.g. in interviews with US-Americans or by US-American journalists).

Table 10. *Impact on/upon our business* vs. *impact the nation's economy* (excluding passives)

	N	PO *(on, upon)*	%	DO (∅)	%	+ ADV	% PO
BNC	58	37 (0.37 pmw)	**63.8**	21 (0.21 pmw)	36.2	6	**100**
COCA 1990–1994	348	71 (0.68 pmw)	20.4	277 (2.66 pmw)	**79.6**	11	**100**
COCA 2015	266	11 (0.55 pmw)	4.1	255 (12.78 pmw)	**95.9**	2	**50**

Similar to what has been observed for the verb *graduate*, an intervening adverbial almost invariably triggers the PO variant, see (11) and (12)[4].

(11) The crisis in the Gulf is **impacting heavily on the economy of the United States** … (COCA, 1990, *CNN News*)

(12) Compulsory water metering will **impact most heavily on the poor, the sick and the elderly**. (BNC, HDT 343)

The data from the COHA given in Table 11 support what the findings from the COCA suggest: the preposition-less variant has become the preferred option in AmE in the course of the 20th century.

Table 11. *Impact* used with the PO and DO variant in the COHA

corpus	corpus size	*N*	PO *(on, upon)*	%	DO (∅)	%
COHA 1900s–1940s	119,386,620	0	0 (0 pmw)		0 (0 pmw)	
COHA 1950s	24,398,180	0	0 (0 pmw)		0 (0 pmw)	
COHA 1960s	23,927,982	1	1 (0 pmw)	**100**	0 (0 pmw)	
COHA 1970s	23,769,305	5	0 (0 pmw)	0	5 (0.21 pmw)	**100**
COHA 1980s	25,178,952	9	3 (0.12 pmw)	33.3	6 (0.24 pmw)	**66.6**
COHA 1990s	27,877,340	54	5 (0.18 pmw)	9.3	49 (1.76 pmw)	**90.7**
COHA 2000s	29,479,451	83	3 (0.10 pmw)	3.6	80 (2.71 pmw)	**96.4**

The question is whether BrE has followed this AmE trend. Because of the sheer size of the GloWbE it was impossible to analyse all forms of the verb exhaustively.

4. There was one counterexample in the COCA: *After controlling for all other independent variables, the type of developmental course taken by a student impacted significantly his or her likelihood of being retained*. (COCA, 2015, *Delta Kappa Gamma Bulletin*)

Thus, I checked all occurrences of *impact* when used as a verb and when directly followed by the preposition *on* or *upon*. In addition, all cases with an intervening adverb as in *The tragedy of losing his mother impacted profoundly on his life and work* (GB G, indepent.co.uk) were also extracted.

Table 12 shows that in current BrE used on the web as represented by the GloWbE only approximately one quarter of all occurrences of *impact* are still used with *on* or *upon*. This is clearly a change when compared to the BNC data that showed the PO variant as the dominant one (63.8%). In AmE, similar to what the COCA and COHA data showed, the PO variant with *on*/*upon* has become marginal in the early 21st century.

Table 12. *Impact* followed by *on*/*upon* in the GB and US components of the GloWbE

corpus	corpus size	IMPACT	+ *on/upon*	+ ADV	% *on/upon*
GloWbE GB	387,615,074	3830	705/100	96	**23.5**
GloWbE US	386,809,355	5860	85/16	10	**1.9**

4.3 *Shop*

The corpus data confirm the dictionary information in that the transitive use of *shop* is typical of AmE usage but hardly attested in BrE (see Table 13).

Table 13. *Shop in/at the supermarket* vs. *shop the farmers' market*

	total count	PO *(in, at)*	%	DO (∅)	%
BNC	95	94 (1 pmw)	**98.9**	1 (0.02 pmw)	1.1
COCA 1990–1994	200	190 (1.83 pmw)	**95**	10 (0.1 pmw)	5

The degree of variation attested in the COCA is admittedly low, but the long-term diachronic development in AmE over the 20th and early 21st centuries suggests that the transitive variant is gaining ground, see Table 14. However, this does not necessarily mean that this development is at the expense of the intransitive pattern. The DO pattern simply emerges as a new variant.

A comparison of recent web data of BrE and AmE presented in Table 15 shows that still very few transitive uses for *shop* can be found in BrE (a mere twelve tokens, suggesting that this pattern has not yet diffused into BrE), while the frequency of occurrence for AmE resembles the one found in the 2010–2015 sections of the COCA (0.22 vs. 0.20 pmw).

Table 14. *Shop* used with the DO variant (any form of the verb followed directly by a noun, determiner, pronoun, adjective or numeral)

corpus	corpus size	DO	DO pmw
COHA 1900s–1960s	167,712,782	1	0.006
COHA 1970s	23,769,305	4	0.17
COHA 1980s	25,178,952	11	0.44
COHA 1990s	27,877,340	15	0.54
COHA 2000s	29,479,451	8	0.27
COCA 1990s	207,446,322	21	0.10
COCA 2000s	204,967,558	39	0.19
COCA 2010–2015	121,375,052	24	0.20

Table 15. *Shop* used in the GB and US sections of the GloWbE (DO variant retrieved by targeting any form of the verb followed directly by a noun, determiner, pronoun, adjective or numeral)

corpus	corpus size	SHOP	+ *at/in*	% *at/in*	DO	DO pmw
GloWbE GB	387,615,074	9220	1442	15.6	12	0.03
GloWbE US	386,809,355	9091	1624	17.9	87	0.22

5. Discussion and conclusion

Based on previous studies that examined different patterns of differential diachronic developments in BrE and AmE, and a critical evaluation of an earlier typology discussed by Marckwardt and Quirk (1964), Hundt (2009: 32f.) proposed a more complex typology of differential grammatical change in BrE and AmE. She identified six patterns of grammatical change:[5]

a) 'true' colonial lag ('extraterritorial conservatism'),[6]
b) 'true' extraterritorial innovation,

5. See also Mair (2002: 110) for an overview of eight possible constellations of grammatical differences and change between BrE and AmE.

6. Hundt is highly critical of the general use of the concept of 'colonial lag' that is commonly used to refer to any apparent retention of mother-country (i.e. British) usage in present-day AmE. She thus suggests the term 'extraterritorial conservatism' as "a more neutral term that includes both colonial and post-colonial language use and avoids the negative implications of 'lag' " (Hundt 2009: 32).

c) divergence from the common ancestor,
d) parallel developments,
e) resurrection or revival of an older feature, either in the extraterritorial variety or the original homeland,
f) 'kick-down development', i.e. AmE starts out as more conservative but overtakes BrE as the change gains momentum.

Considering this typology, the following conclusions can be drawn for the three verbs under study here. For *graduate*, the data show a complex complementation profile with preferences for patterns shifting between BrE and AmE. For Patterns 1 and 3, no BrE equivalent is attested in the research literature and the BNC data do not suggest otherwise. Thus, these uses are most probably cases of extraterritorial innovation in AmE (see also Galinsky 1952: 165, 194). The web data from the GloWbE indicate that Pattern 1 (*graduate university*) may slowly find its way into BrE usage.

For *impact*, the data suggest a steady diachronic development towards a preference for the DO variant in AmE, while the PO variant with *on/upon* is still the majority choice in the BNC. An early transitive use of *impact* is not attested in the research literature (Galinsky 1952, Kirchner 1959). The corpus data thus point towards another case of extraterritorial innovation in AmE, possibly initiated by semantic or structural analogy to other verbs. BrE has followed this trend in the early 21st century with the share of the PO variant dropping from 63.8 to 23.5%. Such "Americanization", i.e. the influence of north American habits of expression and behaviour on BrE (and other national varieties), has been observed for other changes in verbal complementation, e.g. in the spread of the bare infinitive with *help* (see Mair 2002).

As for *shop*, an early transitive use is not included in the long lists drawn up by Kirchner (1959, 1972). As the date of first attestation of the transitive use of *shop* the OED quotes the first edition of Galinsky (1974) published in 1957 that provides the example of a commercial ad in Minneapolis ("Shop the store that gives you more"). The COHA, COCA and GloWbE data suggest that it is again a case of extraterritorial innovation that is gaining ground. More recent BrE data suggest that this use has not (yet) made its way into BrE in the course of the early 21st century. Finally, for both *graduate* and *impact* the data show that in both varieties there is no variation between the PO and DO variant if an adverbial intervenes. This could be caused by the increased structural complexity of the construction as captured by the Complexity Principle (see Rohdenburg 2009: 200).

To summarise and conclude, in view of the fact that with the verbs under study the transitive uses constitute cases of extraterritorial innovation, it appears that prepositions are semantically redundant, not adding significantly to the meaning

of the verb. This chapter thus provides further empirical support for the hypothesis that AmE grammar "shows a more marked tendency to dispense with function words that are semantically redundant and grammatically omissible", and that this "trend towards grammatical economy ties together an array of otherwise unrelated phenomena in the complementation system" (Rohdenburg & Schlüter 2009: 6).

References

Algeo, John. 1988. British and American grammatical differences. *International Journal of Lexicography* 1(1): 1–31. doi:10.1093/ijl/1.1.1

Algeo, John. 2006. *British or American English? A Handbook of Word and Grammar Patterns*. Cambridge: CUP. doi:10.1017/CBO9780511607240

Burnard, Lou. 2007. *Reference Guide for the British National Corpus (XML Edition)*. <http://www.natcorp.ox.ac.uk/docs/URG/> (12 August 2016).

Corpus Resource Database (CoRD). 2016. <http://www.helsinki.fi/varieng/CoRD/index.html>

Davies, Mark. 2004-. BYU-BNC. Based on the *British National Corpus* from Oxford University Press. <http://corpus.byu.edu/bnc>

Davies, Mark. 2008-. *The Corpus of Contemporary American English*, <http://corpus.byu.edu/coca>

Davies, Mark. 2010-. *The Corpus of Historical American English*, <http://corpus.byu.edu/coha>

Davies, Mark. 2013-. *Corpus of Global Web-Based English: 1.9 Billion Words from Speakers in 20 Countries*. <http://corpus.byu.edu/glowbe>

Davies, Mark & Fuchs, Robert. 2015. Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-based English Corpus (GloWbE). *English World-Wide* 36(1): 1–28. doi:10.1075/eww.36.1.01dav

Galinsky, Hans. 1952. *Die Sprache des Amerikaners, Band II: Wortschatz und Wortbildung - Syntax und Flexion*. Heidelberg: Kerle.

Galinsky, Hans. 1974. *Amerikanisches und Britisches Englisch. Einführung in Einheit und Verschiedenheit einer Weltsprache*, 3rd edn. München: Hueber.

Herbst, Thomas, Heath, David, Roe, Ian F. & Götz, Dieter (eds). 2004. *A Valency Dictionary of English*. Berlin: Mouton de Gruyter. doi:10.1515/9783110892581

Hundt, Marianne. 1998. *New Zealand English Grammar - Fact or Fiction? A Corpus-Based Study in Morpho-Syntactic Variation* [Varieties of English around the World G23]. Amsterdam: John Benjamins. doi:10.1075/veaw.g23

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 13–37. Cambridge: CUP. doi:10.1017/CBO9780511551970.002

Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles*, Part III: *Syntax*, Vol. 2. London: Allen & Unwin and Copenhagen: Munksgaard.

Kirchner, Gustav. 1955. Direct transitivation. *English Studies* 36: 15–23. doi:10.1080/00138385508596940

Kirchner, Gustav. 1959. Zur transitiven und intransitiven Verwendung des englischen Verbums. *Zeitschrift für Anglistik und Amerikanistik* 7(4): 342–399.

Kirchner, Gustav. 1972. *Die syntaktischen Eigentümlichkeiten des amerikanischen Englisch*, Band 2. München: Hueber.

Longman Dictionary of Contemporary English. Online version. <http://www.ldoceonline.com> (February 2017).

Mair, Christian. 2002. Three changing patterns of verb complementation in Late Modern English: a real-time study based on matching text corpora. *English Language and Linguistics* 6(1): 105–131. doi:10.1017/S1360674302001065

Marckwardt, Albert H. & Quirk, Randolph. 1964. *A Common Language: British and American English*. London & Washington DC: The British Broadcasting Corporation & The United States Government.

Merriam-Webster's Collegiate Dictionary. Online version <http://www.m-w.com> (February 2017).

Merriam-Webster's Learner's Dictionary.<http://www.learnersdictionary.com> (February 2017).

Oxford English Dictionary, 3rd ed. Revised and updated online version <http://www.oed.com> (February 2017).

Rohdenburg, Günter. 2009. Nominal complements. In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 194–211. Cambridge: CUP. doi:10.1017/CBO9780511551970.011

Rohdenburg, Günter & Schlüter, Julia. 2009. Introduction. In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 1–12. Cambridge: CUP. doi:10.1017/CBO9780511551970.001

Xiao, Richard. 2008. Well-known and influential corpora. In *Corpus Linguistics: An International Handbook*, Vol. 1, Anke Lüdeling & Merja Kytö (eds), 383–457. Berlin: Mouton de Gruyter.

PART 3

The emergence of new patterns

I would like to request for your attention

On the diachrony of prepositional verbs in Singapore English

Sebastian Hoffmann
Universität Trier

This paper employs a diachronic corpus of newspaper data to investigate the use of a range of semantically redundant verb-particle combinations (e.g. *to request for*) in Singapore English since the 1950s. The findings are interpreted as evidence for possible processes of innovation and structural nativization, as outlined in Schneider's (2003, 2007) dynamic model of post-colonial Englishes. The paper also touches on the motivations for speakers to employ such structures (e.g. analogy with existing forms, transparency) and shows that the results of parallel developments can be observed in other New – but not in native – Englishes world-wide, thus suggesting that we are dealing with strategies that are specific to ESL (English as a second language) contexts.

Keywords: prepositional verbs; corpus linguistics; diachrony; post-colonial English; Singapore English; structural nativization

1. Introduction

The past few decades have witnessed a tremendous growth in the study of Englishes worldwide. Apart from detailed – and often corpus-based – descriptions of varietal differences between native Englishes (see e.g. Leech et al. 2009 for British vs. American English, Fritz 2007 for Australian English, and Hundt 1998 for New Zealand English, to name but a few examples), researchers have also increasingly focused on the investigation of institutionalised second-language varieties (ESL) such as Indian English (e.g. Sedlatschek 2009; Schilk 2011), Jamaican English (e.g. Hinrichs 2006) and Hong Kong English (e.g. papers in Bolton 2002). Even lesser-known varieties such as St. Helenian English or Tristan da Cunha English – the latter with fewer than three hundred speakers – have by now received ample scholarly attention (see Schreier 2003, 2008).

DOI 10.1075/slcs.195.10hof

From a theoretical perspective, this geographical expansion of varietal interests has gone hand-in-hand with the development of models of English worldwide. In addition to classification systems such as the influential three-circle model by Kachru (1985) or the five-category model proposed by Gupta (1997), these include models of the developmental phases of New Englishes. Well-known examples are Kachru's (1982) three-phase model, involving the steps of "non-recognition", "co-existence" and "recognition" and Moag's (1982) "life-cycle of non-native Englishes", with five processes that are (potentially) undergone by new varieties. The most influential of these theoretical approaches has no doubt been Schneider's (2003, 2007) dynamic model of the evolution of New Englishes, which also lists five phases, and which accounts for the complex interplay between socio-political developments, identity construction, the sociolinguistics of contact and structural effects on linguistic forms.

The field of World Englishes continues to be a thriving enterprise, with interests in recent years shifting to – for example – the investigation of parallels between ESL and English as a Lingua Franca (ELF) or English as a Foreign Language (EFL) contexts (see e.g. papers in Mukherjee & Hundt 2011 and Schneider 2012), and with methodological sophistication continually improving in order to capture subtle differences and parallels between varieties that would not emerge from an evaluation of simple descriptive statistics (see e.g. Mukherjee & Gries 2009, which studies collostructional nativization in New Englishes).

However, while synchronic studies of (present-day) English worldwide clearly abound, hardly any diachronic studies exist that chart the development of ESL varieties in real time. The reason for this is very simple: There is as yet no suitable database/corpus that would allow researchers to trace linguistic changes on an appropriately large scale. As a consequence, scholars have so far needed to resort to an approach that resembles the apparent-time approach prominent in sociolinguistic studies, interpreting differences between various ESL varieties and their input variety – typically British English – as indicative of a placement in the different phases of Schneider's model (see e.g. Mukherjee & Hoffmann 2006).[1] As a corollary of this, no studies exist that thoroughly test whether the development of new varieties does indeed follow the models proposed.

1. There are some notable exceptions, however. See for example Borlongan & Dita (2015) on the investigation of changes in Philippine English across a 30-year period. Also, a number of projects are underway to create suitable diachronic sources for the study of ESL varieties; see e.g. Biewer et al. (2014) and Brato (2014).

A number of methodological problems arise as a result of the lack of diachronic data. These include the difficulties in assessing

- the distinction between learner language phenomena (errors) and genuine nativized regionalisms (particularly as regards low-frequency features)
- the role of first-language interference
- the distinction between fossilised features of earlier stages of the input variety and real innovations (cf. the concept of "colonial lag", Görlach 1987)

As a case in point, consider examples (1) – taken from an Indian broadsheet daily newspaper – and (2):[2]

(1) The employee is also required to *inform* the appointing authority the amount of monthly instalment… (*The Statesman* 1.8.2004)

(2) I have often the satisfaction of hearing the publican, the baker, and sometimes even the parish-clerk, petitioning my housekeeper … to *inform* him the exact time by Master Humphrey's clock.
(Charles Dickens: *Master Humphrey's Clock*, 1840)

Without suitable diachronic data, the ditransitive use of *inform* shown in example (1) could be interpreted to be an innovation in Indian English. As example (2) shows, however, this use was present in earlier British English (BrE), where it is now no longer considered grammatical. But even awareness of such an earlier BrE use will not allow the researcher to claim with any certainty that we are indeed dealing with an archaic remnant of colonial English; only a diachronic description revealing constant use of this structure throughout the period since colonisation would make such a claim fully valid.

This paper is a first step towards a thorough diachronic description of one ESL variety, namely Singapore English (SingE). As such, it is intended to be a pilot study, involving the investigation of one linguistic feature – prepositional verbs – on the basis of only a single text type, viz. newspaper English. I will be interpreting the findings against the background of existing models of the evolution of postcolonial Englishes, in particular Schneider's (2003, 2007) model. I will also touch on some of the methodological difficulties in undertaking this type of study and call for further diachronic studies to be carried out – and further diachronic resources to be created – to form a basis for a better understanding of the processes at work.

2. These examples were also discussed in Mukherjee & Hoffmann (2006).

2. Background

2.1 Singapore English

Singapore English is the variety of English spoken – and written, see Section 2.2 – by the approximately 5 million inhabitants of the Southeast-Asian island state of Singapore. A full description of its history and its present-day features is beyond the scope of this paper; readers who wish to approach the topic in a more extensive fashion are instead referred to Leimgruber (2013a). For a brief but highly informative overview of the linguistic ecology of Singapore, see Gupta (2008). The following points will have to suffice in the context of the present study:

- Like many other New Englishes, SingE is heavily influenced by contact with a range of different first languages. According to the 2010 census (Singapore Department of Statistics 2010: viii), 74.1% of all residents are of Chinese ethnicity, with Malays and Indians constituting relatively small minorities of 13.4% and 9.2% of all residents, respectively.
- The Singaporean government has strong views on language policy and takes a very active stand in influencing the linguistic landscape of the country. It has vigorously pursued a bilingual education policy, with English being the language of instruction at all levels. Although a number of Chinese and Indian dialects/languages are spoken in Singapore, Tamil and particularly Mandarin are on the rise due to their status as "official mother tongues" (see e.g. Stroud & Wee 2010). These mother tongues, along with Malay, are supposed to be used for intra-ethnic communication; furthermore, they are intended to carry the desirable values of the Asian cultures they represent.
- The government's stance on English is heavily exonormative, discouraging the establishment of affective uses of the language that would reflect a local cultural identity. English is to be used for intra-ethnic and external communication exclusively. However, in contrast to its success in fostering the use of Mandarin among the population, the government has so far squarely failed in their quest to eradicate a colloquial, spoken variant of SingE, commonly referred to as Singlish. At the same time, the Speak Good English Movement (SGEM), launched in 2000 to "encourage Singaporeans to speak grammatically correct English that is universally understood" (http://goodenglish.org.sg), has had some impact on reducing the use of a number of perceived non-standard forms, particularly in the educational context.[3]

3. For a detailed description of the SGEM and a critical evaluation of its practices and successes, see Bruthiaux (2010) and Wee (2014).

– SingE is said to be the ESL variety that has progressed furthest in Schneider's (2003, 2007) model. While Schneider (2007: 160) himself places it in Phase IV ("endonormative stabilisation"), there are good reasons to claim that it has advanced even further; in fact, we are currently witnessing a process of language shift, with English becoming a native language for an increasing number of younger Singaporeans.

Given its highly advanced developmental status, SingE is particularly well-suited for an investigation into its diachronic development.

2.2 Singapore standard English?

Considering the intrinsically spoken nature of Singlish, a diachronic account of its development is virtually impossible to achieve, absent the discovery of a large repository of recordings of informal language use in one of the Singaporean archives. The focus of this paper will thus lie on standard written language only. However, the concept of (International) Standard English is a controversial topic. While some scholars emphasise the pluricentric nature of the language (see e.g. Leitner 1992), others stress the fact that these different Standard Englishes in fact share an overwhelming number of their features and that it does not therefore make much sense to talk of local standards; instead, it is claimed that there is a single, global Standard English that is tolerant to limited local variation on some levels of description (e.g. pronunciation and lexis).[4] This is also the view of Gupta (2010), who discards the concept of Singapore Standard English, and thus revises her own views presented in Gupta (1986), where she described a whole range of features whose endonormatively stabilised status would give them a diagnostic value for establishing a local standard.

In my view, there are good reasons to work with the concept of Standard Singapore English. While I agree that most overt features of Standard Englishes worldwide are indeed shared, this is not the whole picture. Endonormativity is expressed not only via the acceptance and codification of obvious localisms in dictionaries and grammars, but it also shows itself more subtly, namely via different distributions of the shared features of so-called Standard English across the different local probabilistic grammars. Differences between Standard Englishes are thus more quantitative and less qualitative. Corpus-linguistic analyses are well-suited to uncover such tendencies, which in their sum can then be used to describe a local standard.

4. See also the concept of a "common core" used in Quirk et al. (1985: 15).

2.2 Prepositional verbs

Prepositional verbs are a subset of multi-word verbs – sometimes also referred to as particle verbs – that behave as a single unit even though they are spelled as two orthographic words. A typical example is shown in (3), while (4) and (5) display examples of a phrasal verb and a phrasal-prepositional verb, respectively. For a full overview of their features, see Quirk et al. (1985: 1150–1168).

(3) She *looked at* the picture. (prepositional verb)

(4) He *turned off* the light. (phrasal verb)

(5) I *look forward to* seeing you next week. (phrasal-prepositional verb)

The most prominent distinction between phrasal and prepositional verbs is that the former allows the particle to be moved to a position after the following noun phrase.[5] Consider (3a) and (4a) below:

(3a) *She *looked* the picture *at*.

(4a) He *turned* the light *off*.

Prepositional verbs can be followed by two noun phrases (e.g. *She invested her inheritance in index funds*), in which case the first is the direct object and the second is the prepositional object. The current paper is only concerned with prepositional verbs that do not have direct objects; in other words, with the exception of optional adverbials, the lexical element of the multi-word verb is always adjacent to the particle.

Prepositional verbs in World Englishes have received considerable scholarly attention. A prominent example is Schneider's (2004) study of particle verbs, i.e. all three types shown in examples (3) to (5). Although findings are somewhat tentative due to the relatively small numbers retrieved from five ICE-corpora (including ICE Singapore), he discovers systematic differences in overall frequencies between the varieties studied, which he interprets as evidence of structural nativization.

Interest in prepositional verbs is further motivated by the fact that ESL varieties tend to exhibit prepositional verb usages that would not be considered grammatical in native Englishes such as BrE or AmE; see e.g. Bamgbọṣe (1982) for Nigerian English, Gonzales (1983) for Philippine English and Tan (2013) for Malaysian English. As described in Deterding (2007) and Tan (2016), this is also the case for SingE. Typical examples are displayed in (6) to (9), all taken from Tan (2016):

5. Another criterion is stress in the corresponding passive forms, with stress in prepositional verbs normally occurring on the lexical element of the multi-word verb. See Quirk et al. (1985: 1157).

(6) Assoc Prof Simon Tay of the NUS Faculty of Law … *discussed about* the emerging character of the new Chinese leadership …
('Print media coverage', National University of Singapore website, May 2013)

(7) 'The largest possible role for research is …,' said Mr Lai, while *emphasising on* the numerous opportunities afforded to businesses in such contexts.
(*The Business Times*, 6 March 2012)

(8) He believes an easier or more direct way would be to restrict the number of foreign workers *entering into* the country.
('Risks and opportunities for the construction industry', Singapore Management University website, 3 August 2011)

(9) If I cannot complete the formalities within the 2-month period, can I *request for* an extension?
('FAQs', The Immigration and Checkpoints Authority of Singapore website, 2 September 2013)

In all of these sentences, speakers of BrE or AmE would not employ a prepositional verb but opt for a version without a particle. From a semantic point of view, the particle does not provide further relevant information; the verb would offer sufficient information in itself and the particle is thus semantically redundant.

It is highly instructive to look at the sources of examples (6) to (9). As Tan (2016: 72) notes, these multi-word verbal combinations "are being used by some of the strongest proponents of the standard language ideology in Singapore. This seems to suggest some degree of institutionalization – that is to say that some of these [prepositional verbs] may have become permanent variants in SgE."

Prepositional verbs have also received attention from research into the differences and – more importantly – commonalities of learner English and second-language varieties; see e.g. Nesselhauf (2009) and Schneider & Gilquin (2016). I will return to these possible parallels in Section 6 below.

The present study aims to test whether a diachronic analysis of the use of prepositional verbs of the type shown in examples (6) to (9) can provide evidence for the processes of innovation and structural nativization in SingE. I will also touch on the motivations behind the establishment of such innovative forms – that is, if they are indeed determined to be innovations rather than archaic remnants of earlier stages of the input variety.

3. Data

The bulk of my data stem from an electronic collection of 158 million words of newspaper data representing seven individual years (1951, 1961, 1971, 1981, 1991, 2001 and 2011) of the *Straits Times*, the main English-language daily newspaper

in Singapore.[6] The *Straits Times* is a Singaporean broadsheet paper with a history going back to 1845. It currently has a circulation of approximately 350,000 copies and its political orientation is ostensibly parallel to the views of the People's Action Party (PAP), which has ruled Singapore since 1959.

Although all seven years were provided in electronic format, there are some important differences between the years 1951–1981 and the three latest years: While the latter years were from the outset keyed in electronically and thus fully represent the original source of the newspaper texts, the earlier four years are based on paper copies that were processed with optical character recognition (OCR) software. Unfortunately, this introduced a whole range of errors in the data, which somewhat reduces the quality of the data. As a case in point, consider examples (10) and (11):

(10) STAMPED a TvDlni SDeed 4swom SELF ADDRESSED envelope iypingspeeo.43wprn (Slie 23 cm × 10- cm If you are n.ri.nr." *requesting for* application forms 1 by post). (*Straits Times*, 6.1.1981)

(11) Thai Ambassador to sia is san Vejjajiva said that they had *requested for* high yielding rubber, coconut, oil palm and cocoa seedlings. (*Straits Times*, 12.2.1981)

While the text in (10) is seriously affected (possibly as a result of graphic elements that were wrongly identified as lexical material), example (11) displays much less distortion as a result of the OCR process. In the latter case, *sia* should read *Malaysia*, and the ambassador's first name should be given as *Nissai*.[7] Fortunately, the large majority of texts do not exhibit the difficulties shown in (10), and even if the texts searched contain some errors, it is relatively unlikely that they will affect the search terms used for retrieving prepositional verbs themselves.

A second important difference relates to the fact that the data for the years 1991–2011 does not cover every single word printed in a daily edition of the newspaper but covers only those texts that could clearly be identified as having been written/edited locally. This leads to a slight imbalance between the two parts, which again reduces comparability. Also, as a result of this I have more data for each year in the earlier data; this is particularly the case for the year 1981, which alone accounts for about 55 million words in the corpus.

6. The newspaper texts were licensed in 2012 from the Singapore Press Holdings (SPH) for a period of three years, after which only 10 per cent of the texts can be used without any time limit. The data for this project were retrieved before the license for full access expired.

7. Such errors may be the result of bad photocopies, with margins of texts being rendered as black patches rather than actual words.

The primary data for the paper are complemented by two additional resources: GloWbE and the NOW corpus. GloWbE is a corpus of global web-based English comprising about 1.9 billion words from 20 different countries (see Davies & Fuchs 2015); the Singaporean component of the corpus amounts to approximately 43 million words. About 60 per cent of GloWbE reflects less formal, more speech-like uses as found in blogs and discussion boards, while about 40 per cent contains data from webpages that on the whole represent somewhat more formal language usage.[8] NOW stands for News on the Web, and the corpus currently comprises about 3.8 billion words of newspaper data from 2010 onwards, covering the same 20 countries represented in GloWbE. The NOW corpus is a monitor corpus that is being updated on a daily basis and thus allows researchers to investigate large amounts of data (almost) as they are being produced.

Both the *Straits Times* and the GloWbE data were available to me in "raw text" format and were thus searchable with Perl scripts, while the NOW corpus could only be accessed via the web interface provided (http://corpus.byu.edu/now/). The method of retrieving prepositional verbs will now be described in Section 4.

4. Retrieving prepositional verbs

Given the pilot nature of the present study, I did not aim at a fully-comprehensive analysis of all prepositional verbs in the data. However, in order to draw meaningful conclusions about the status of verb-particle combinations in SingE that would be unusual or impossible in native Englishes, I had to retrieve a sufficient number of relevant examples of the constructions at hand. For this purpose, I compiled frequency lists of potential nativized/nativizing prepositional verbs in the Singaporean component of GloWbE and compared the results to frequencies of the same combinations obtained from the two major native-speaker components of the corpus (i.e. data from the US and Great Britain, amounting to 774 million words) by calculating a factor of their frequencies per million words in the two sets of data. I then extracted for further analysis those combinations in the *Straits*

8. It is worth noting here that a considerable chunk of the data in both the more and the less formal components of the corpus consists of comments provided by readers of the webpages rather than by their authors. My current estimate is that these comments amount to about 200 million words, or more than 10 per cent of the whole corpus. Although the Google-based categorization method to identify web sites by country appears to be very robust (see Davies & Fuchs 2015: 4–5), some level of distortion regarding the source of linguistic forms found in the corpus may of course nevertheless be introduced via readers from other countries who are providing comments.

Times that 1) only existed in the SingE data or 2) exhibited particularly great differences in their overall frequencies.[9]

GloWbE contains part-of-speech (POS) information (as provided by the CLAWS tagger, see Garside 1987), which allowed me to restrict the retrieval to POS-tag combinations that represent potentially relevant items. However, the area of prepositional verbs is prone to tagging errors – in particular with respect to combinations that would be unusual or impossible in native Englishes. This is demonstrated in example (12), where the verb *requests* is tagged as plural noun.

(12) Data_NN is_VBZ translated_VVN only_RR when_CS a_AT1 user_NN1 *requests_NN2 for_IF* a_AT1 specific_JJ piece_NN1 of_IO information_ NN1,_PUN "_PUQ said_VVD Mr_NNB Ang_NP1._PUN

In order to reduce the impact of such tagging errors, a number of strategies were implemented. For example, I also compiled frequency lists of any word followed by any of the particles/prepositions listed by Quirk et al. (1985: 1151). This retrieved far too many types to make a manual analysis possible, but it allowed me to check whether there are relevant high-frequency combinations typical of SingE that are not found via a purely tag-based approach – which was in fact not the case. A second strategy was to temporarily restrict retrieval to forms with verbal morphology only (i.e. words ending in *-ing* and *-ed*), thus avoiding the retrieval of nominal homographs of potential items of interest, such as *(the) request for*. While a claim of completeness would certainly be unwarranted, I am confident that I have retrieved the forms with the most prominent frequency differences between SingE and native varieties in GloWbE. Table 1 lists the 14 forms retrieved via the method just described.[10]

---

9. There are more sophisticated methods of retrieving potentially relevant combinations; see e.g. Schneider and Gilquin (2016) who calculate collocational strengths of automatically retrieved verb-preposition/particle pairs in ENL, ESL and learner data that are then sorted according to a collocation ratio (see also Schneider & Zipp 2013), thus detecting combinations that are particularly overused in non-native contexts. The frequency-based approach taken in the present paper is unlikely to miss prominent examples of nativised prepositional verbs, though.

10. The method also retrieved a number of potentially nativized phrasal verbs such as *to end off*, *to list down* and *to voice out*. However, these are outside the scope of the present paper. Also, my list contains a number of combinations that do not feature in Schneider & Gilquin's (2016) tables; on the other hand, they include *to study about*, which was not found with a sufficiently large frequency difference in my data to warrant inclusion. As mentioned above, it is not the aim of the present study to conduct a fully-comprehensive analysis of all prepositional verbs. The prepositional verb *to enter into* was not determined to be particularly frequent in SingE

Table 1. Potential nativized prepositional verbs in SingE

to await for	*to leverage on*
to comprise of	*to mention about*
to crave for	*to request for*
to discuss about	*to research on*
to emphasise on	*to share about*
to enter into	*to target at*
to investigate on	*to test on*

It must be noted that of course not all instances of the forms shown in Table 1 actually function as prepositional verbs. As a case in point, consider examples (13) and (14), only the first of which is a relevant instance of a prepositional verb:

(13) [I] know its gonna be a huge play of emotions that *awaits for* us tonight. (GloWbE, <http://joonni.com/2012/10/28/on-the-eve-of-the-final-week-of-faith/>)

(14) The cost of travel for your family from the US, UK, or elsewhere may not be cheap, but once you arrive the holiday of your dreams *awaits for* a fraction of the cost back home. (GloWbE, <http://www.thailandfamilyvacation.com/>)

For the discussion presented in Section 5, all forms presented in Table 1 were searched in the *Straits Times* data in order to detect possible diachronic developments. Naturally, all instances that were retrieved were manually checked for their relevance as prepositional verbs.

5. Results

As previously indicated, the aim of this paper is to trace the diachronic development of a number of prepositional verbs in the *Straits Times* data and to use these findings as evidence for possible processes of innovation and structural nativization in SingE. In order to do so, the forms are required to exhibit a minimum number of hits. Some of the items displayed in Table 1 could therefore not be subjected to further analysis. Secondly, it does not make much sense to investigate forms that for language-external reasons cannot be expected to have been used in earlier decades. This restriction applies in the context of *to share about*, whose

by the comparison of GloWbE components. Its inclusion is instead motivated by its explicit mention in Tan (2016) – see below.

frequencies in GloWbE single it out as a decidedly Singaporean feature by being markedly more frequent in the Singaporean component than in any of the other native or non-native varieties covered by the corpus. Its prepositional verb usage is shown in example (15):

(15) Would you like to *share about* your current emotional state and what brought you here? (*Straits Times*, 1.10.2011)

No uses of this type prior to 2011 can be found in my data, and it is highly likely that this prepositional verb will have been brought about by the extended use of the verb *to share* on social media, most prominently Facebook, where it refers to making a story or a link available to a user's social circle. Similar uses can also be found in native varieties of English (including in newspaper texts), albeit at lower frequencies.[11] The data thus show that speakers of SingE may be particularly likely to adopt new prepositional verbs into their inventory, but given the lack of diachronic depth and the fact that *to share about* is used in native varieties of English, it will not be further discussed here. In what follows, I will particularly be concentrating on four of the items shown in Table 1, namely *to enter into*, *to await for*, *to request for* and *to leverage on*.

5.1 The prepositional verbs *to enter into* and *to await for*

As indicated in footnote 10 above, *to enter into* was included in the field of prepositional verbs for further analysis even though its relative frequency does not vary greatly between native varieties of English and SingE. However, Tan (2016: 73) explicitly lists it among her nativized multi-word verb combinations and it thus warrants further attention.

As a matter of fact, the prepositional verb *to enter into* is common in native varieties of English. Typical examples, taken from the British National Corpus (BNC) are shown in (16) to (18):

(16) This is not the place to *enter into* a debate about contemporary sports policy. (BNC, A6Y:1139)

(17) At that stage you can *enter into* a new covenant. (BNC, A01:274)

(18) […] a very sympathetic registrar who managed to *enter into* the spirit of things. (BNC, A7P:929)

11. The frequency of all word-forms of *to share about* is about seven times higher in SingE than in the two native varieties. However, many of the native speaker instances are not relevant prepositional verb uses (e.g. *Jeff, Itay and Ruth from ICAHD, with whom we* shared about *a ton of pizza in a tent in Beit Arabiya […]*, GloWbE, <http://georgeluke.wordpress.com/>).

In all three cases – and in all of the other BNC examples – *to enter into* is used in an abstract context. Other strong collocates than those shown in (16) to (18) include *contract*, *agreement(s)*, *transaction(s)*, *correspondence* and *negotiations*. However, in earlier stages of the language, a much more literal meaning of *to enter into* was predominant; cf. examples (19) to (21) taken from Early English Books Online (EEBO):

(19) And than the kyng *entred into* the towne & rested hym in the castel tyll the towne was set in rule & in gouernance. (EEBO, 1528)

(20) When thou dost pray, *enter into* thy chamber, and shutting fast thy door, pray unto thy heavenly father. (EEBO, 1563)

(21) When thy meate firste *entreth into* thy stomake, the naturall heate beginneth (as it were) to make warre with thy meate vntill that the substance thereof be altered and chaunged […] (EEBO, 1584)

As examples (22) and (23) show, such literal uses can also be found in the *Straits Times*:

(22) YOU *enter into* an apartment and you are not sure if it is the living room or the bedroom that greets you. (*Straits Times*, 21.7.2001)

(23) An estimated 50 to 70 per cent of China's trade is processing trade – that is, products *entering into* China to be assembled before being shipped out again for consumers in other markets. (*Straits Times*, 4.11.2011)

It could thus be hypothesized that we are here dealing with an archaic use of the prepositional verb, preserved perhaps as a result of colonial lag. However, as Table 2 shows, a search in the *Straits Times* data produces a sobering number of hits: of the 1,519 instances of *to enter into* – in all its word-forms – only four represent literal uses of the prepositional verb.

Table 2. Instances of literal *to enter into* in the *Straits Times*

1951	–
1961	–
1971	–
1981	–
1991	–
2001	2
2011	2

In contrast, about 15 per cent of all instances of *to enter into* in the Singapore component of GloWbE exhibit a literal use.

The raw frequencies presented by the *Straits Times* data are thus far too low to draw reliable conclusions. Nevertheless, the information contained in Table 2, combined with the relatively high proportion of literal examples in the web-based data of GloWbE, suggests that a colonial lag scenario is perhaps an unlikely interpretation. Only the availability of larger – and more diverse – sources of diachronic data can confirm this hypothesis.

A somewhat different picture emerges from the data for the prepositional verb *to await for*. Like literal *to enter into*, it represents a use that is now obsolete in native Englishes – see senses 7 and 8 for *await*, v. in the *Oxford English Dictionary* (*OED*), which have illustrative examples from the 17th century. Two typical instances are shown in (24) and (25).

(24) The British High mission here was still *awaiting for* details of the new tariff rates. (*Straits Times*, 10.12.1971)

(25) "With the market still clouded with uncertainties and the downside risks outweighing the upside potential, we would advise investors to stay on the sidelines and *await for* better emerging values," he said. (*Straits Times*, 30.10.2011)

As Table 3 shows, this prepositional verb is found throughout the period covered by the *Straits Times* data, albeit at low frequencies. There is obviously little sense in expressing such low absolute numbers in relative frequencies; this information is simply given to alert readers to the fact that the corpus contains unequal amounts of data per decade.

Table 3. Instances of *to await for* in the *Straits Times*

Year	*n*	frequency pmw
1951	2	0.11
1961	1	0.04
1971	4	0.12
1981	20	0. 3
1991	3	0.26
2001	–	0
2011	1	0.07

In contrast to *to enter into*, we thus have some evidence for a possible retention of an archaic feature of BrE. However, here, too, further data is required to confirm this hypothesis.

5.2 The prepositional verb *to request for*

The prepositional verb *to request for* is one of the most frequent combinations in my data. A typical example of its use was already given in (9) above; three further instances from the *Straits Times* data are shown in (26) to (28):

(26) JOSEPH YANG Singapore: I WOULD like to *request for* an increase in the frequency of Service 154. (*Straits Times*, 20.7.1991)

(27) Customers repeatedly *requested for* fresh produce and more variety, and Cold Storage extended the range to include fresh produce in April 1999. (*Straits Times*, 23.5.2001)

(28) The public is also allowed to *request for* flavours which makes us feel that we have a say in the brands. (*Straits Times*, 13.3.2011)

The frequency of this prepositional verb displays an interesting development over the time period covered by my data. As Figure 1 shows, there are no instances in the year 1951, while there is a steady increase from 1961 to 2001. This development suggests that a process of nativization is indeed taking place.

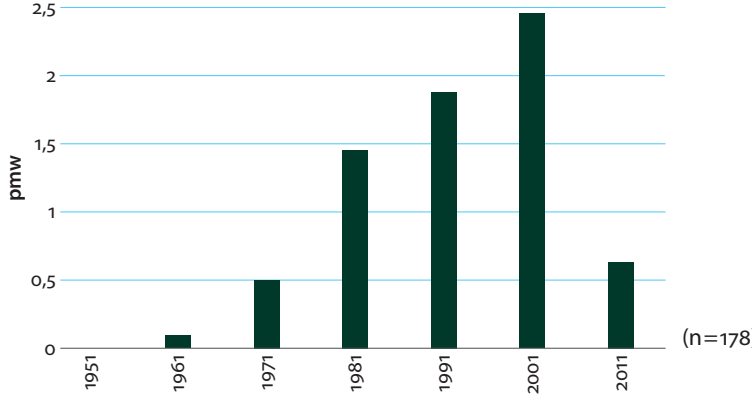

Figure 1. The frequency of the prepositional verb *to request for* in the *Straits Times* data (pmw)

But how do we interpret the sharp decrease in frequency shown in 2011?[12] One can only hypothesize, but my hunch is that we are here dealing with the result of exonormative pressure in an environment that is heavily subject to editorial intervention. I have no proof to substantiate this claim, but I can imagine that *to request*

12. The frequency difference between 2001 and 2011 is highly significant (X^2 = 16.13257, $p < .001$).

for – and perhaps similar "non-standard" prepositional verbs? – is/are now explicitly listed and proscribed in the internal style guide of the *Straits Times*, which has resulted in most of its uses being purged by the editors.

Although *to request for* is one of the more frequent prepositional verbs under investigation, its overall frequencies – n=178, maximum 2.5 instances pmw per decade – are still rather low, which perhaps reduces the reliability of my interpretation considerably. However, the claim that we are indeed dealing with a nativizing construction can be further substantiated by considering the full envelope of variation rather than mere frequencies. In other words, we will be able to assess the development more conclusively by also investigating those instances where *to request for* could have been used but in fact was not.

For this purpose, I retrieved all instances of *request* in the *Straits Times* data and determined 1) whether it was a form of the verb *to request,* and if so, 2) whether an alternative with the preposition/particle *for* would have been possible. In order for this to be the case, the following instances of verbal *request* had to be discarded:

- mandative subjunctives (e.g. *They **requested** that the school keep it open*)
- *to*-infinitives (e.g. *The Chinese side has **requested** to publicise the figures*)
- passives (e.g. *The government does not plan to reintroduce the (English) policy, as **requested** by several parties*).[13]

Given that the lemma *request* has a total of 12,549 in the corpus, I worked with random samples of 400 instances per decade. The numbers given in Table 4 below are thus partially extrapolated results.

Table 4. Instances of verbal *request* with and without *for* in the *Straits Times* data

Year	no *for*		*for*		total
	n	%	*n*	%	*n*
1961	83*	97%	2	3%	85
1971	74*	85%	13	15%	87
1981	296*	77%	89	23%	385
1991	92*	80%	23	20%	115
2001	119*	76%	40	24%	159
2011	130*	92%	11	8%	141

*extrapolated figures based on random subsets of 400 instances per decade

13. This exclusion might be seen as slightly controversial, since it would in theory be possible to say something like *as requested **for** by several parties*. However, I did not find any instance of this type in my data and therefore decided against including passives. The percentages presented in Table 4 would be considerably lower if all passives were counted as possible instances for the preposition/particle *for* to occur.

Table 4 displays the total number of instances of *to request for* per decade as a percentage of all relevant instances of verbal *request*. The data clearly show that the writers opt for the prepositional verb in a considerable proportion of all instances of verbal *request*. In the years 1981, 1991 and 2001, between a fifth and almost a quarter of all possible choices were realised as prepositional verbs. In the year 2011, this drops to only eight per cent, almost returning to the state of the 1960s.

This data strongly suggest that structural nativization is indeed taking place, and that the proportional increase from a low single-digit figure to almost a quarter of all instances nicely corresponds to the projections made by Schneider's (2003, 2007) dynamic model, which postulates Phase III ("structural nativization") to occur between ca. 1945 and ca. 1970s, with Phase IV ("endonormative stabilization") starting in the 1970s. For further details, see Schneider (2007: 155ff.). Obviously, "one swallow does not a summer make", and further diachronic results of this type are necessary to offer more conclusive confirmation of Schneider's model.

5.3 The prepositional verb *to leverage on*

The final item to be discussed is *to leverage on*, which is a prepositional verb that, to the best of my knowledge, has not yet been described in the literature. With 219 instances in my data, it is even more frequent than *to request for*. Before looking at typical examples, it will be helpful to consider the entries for the verb *to leverage* in two major dictionaries, the *OED* and *Merriam-Webster*. According to the former, *to leverage* is an AmE verb meaning "[t]o lever; *spec.* to speculate or cause to speculate financially on borrowed capital expecting profits made to be greater than the interest payable" (*OED*, s.v. *leverage*, v.).[14] In contrast, Merriam-Webster lists two senses:

1. to provide (something, such as a corporation) or supplement (something, such as money) with leverage; *also*: to enhance as if by supplying with financial leverage
2. to use for gain: exploit <shamelessly *leverage* the system to their advantage – Alexander Wolff> (*Merriam-Webster*, s.v. *leverage*, v.)

The meaning of *leverage* is thus clearly related to the financial domain and potentially carries negative connotations – in the latter case presumably also in extended uses.

---

14. The *OED* notes that the entry for *leverage*, v. was first published in 1976 and has not been fully updated yet.

Interestingly, the instances of *to leverage on* in the *Straits Times* typically do not conform to those dictionary entries; representative examples are shown in (29) to (31). The domain of use is clearly not exclusively financial, and the connotations are much more positive than suggested by *Merriam-Webster*. Thus, the meaning could be paraphrased as 'profit from' or 'make good use of'.

(29) Established chipmakers in Singapore and Taiwan can also *leverage on* their supply chain infrastructures, stable government policies and intellectual capital, said Gartner Asia-Pacific semiconductor analyst Tan Kay Yang. (*Straits Times*, 1.10.2001)

(30) It is up to the Government to adapt and *leverage on* social media, he said. (*Straits Times*, 20.10.2011)

(31) Over the years, we have implemented initiatives that *leverage on* information technology to enhance teaching and learning experiences and outcomes. (*Straits Times*, 20.10.2011)

As Figure 2 shows, the use of *to leverage on* does not show a gradual increase of the type seen with *to request for*. Instead, this usage appears quite suddenly; its first attestations in data from the year 2001 are already more common than *to request for* in its most frequent decade (cf. Figure 1 above).

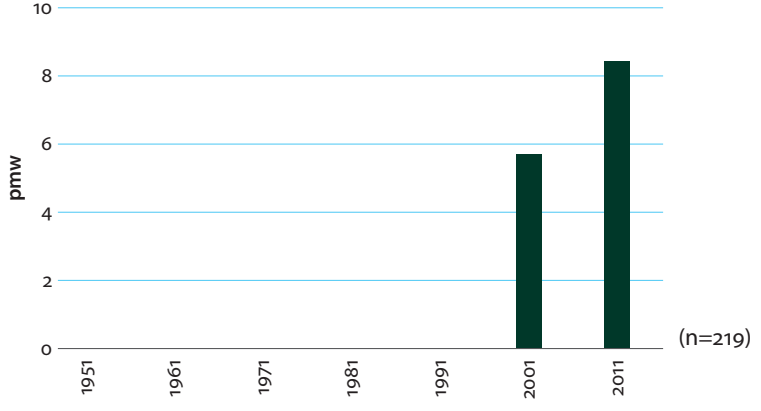

Figure 2. The frequency of *to leverage on* in the *Straits Times* data

In addition, verbal *leverage* is virtually absent from the *Straits Times* data in the 20th century (i.e. 1951–1991), but once it is used, the overwhelming number of instances are found in conjunction with the preposition/particle *on*. In contrast, this type of use is almost completely absent in the native-speaker varieties of GloWbE, suggesting that it is indeed an ESL innovation.

Before providing further quantitative results on the use of *to leverage on*, it is necessary to focus on the motivations behind the development of these types of nativized prepositional verbs. This will then allow me to offer some more general remarks about prepositional verbs in Englishes worldwide and to present an outlook of possible research avenues to develop based on these considerations.

6. Prepositional verbs in new Englishes

Section 5 presented quantitative findings for four prepositional verbs that are found in my Singapore data, but would not be considered acceptable uses in native varieties of English. But why do these forms exist? A number of possible reasons can be proposed, all of which have previously been mentioned in the literature (see in particular Tan 2016):

- retention of obsolete BrE uses
- first language (L1) interference/substrate influence
- formation by analogy
- formation as a result of increased expressiveness/clarity/transparency

With the exception of *to await for*, the data presented here suggest that we are not dealing with the retention of older forms; all the remaining prepositional verbs are thus innovations in an ESL context. In particular, the findings for *to request for* are suggestive of a gradual nativization process, with up to a quarter of all possible instances covered by the new form in more recent decades of the data.

L1 interference – or contact-induced change – is perhaps the strongest candidate for many of the characteristics observed in new Englishes; indeed, a whole range of features of Singlish, the colloquial variant of SingE, can clearly be traced back to forms and structures in Chinese dialects (see e.g. Bao 2009, 2010), and to a lesser extent in Malay (cf. Bao & Wee 1999).[15] However, there is little evidence that this factor is key in influencing the establishment of innovative prepositional verbs in SingE (Siew Imm Tan, personal communication).[16] As also pointed out

15. As Leimgruber (2013b: 5) notes, "the often-named input of Tamil […] has been minimal at best."

16. There are some prepositional verbs in Malay that could theoretically serve as a source, e.g. *berbincang tentang* 'to discuss/talk about', *menyebut tentang* 'mention about' and *terdiri daripada* 'comprise of'; however, these may be modern Malay constructions that have in fact been modelled on English.

by Schneider & Gilquin (2016), analogy is a much more likely candidate, given the obvious parallels between most prepositional verbs shown in Table 1 above and related forms that are frequent in native Englishes, too. Thus, *to request for* may be motivated by the parallel use of the nominal *request for* and *to discuss about* has clear parallels to the (prepositional) verb *to talk about sth*. With *to leverage on*, analogical parallels are slightly more difficult to see. It could however be hypothesised that *take advantage of* and *capitalise on* are sufficiently similar to serve as a model; in the case of the latter option, the preposition/particle is the same one as in *to leverage on*.

The final point on the bullet list given above, i.e. formation as a result of increased expressiveness/clarity/transparency, is closely related to the processes involved in analogical extension as it offers an explanation for why such processes might be particularly likely in an ESL context. As Nesselhauf (2009: 22) notes, prepositional verbs (and other multi-word verbs) are more common in ESL and EFL contexts than in native varieties of English. This may be motivated by an element of transparency and regularisation involved in reducing the structural difference between, for example, *the request for X* and *to request X*, resulting in "iron[ing] out a few of the creases in the fabric of the language" (Deterding 2007: 58). In a similar vein, Tan (2016: 72) sees the processes of group second language acquisition at work (cf. Winford 2003: 235–7), which favour "strategies of simplification and overgeneralisation".

If these processes are at the heart of the development, one would assume that similar results could be found in institutionalised second-language varieties across the globe. As suggested by the interest of scholars like Bamgboṣe (1982) for Nigerian English, Gonzales (1983) for Philippine English and Tan (2013) for Malaysian English, this is indeed the case. In this context, I will briefly return to the prepositional verb *to leverage on*, tapping into the 3.8 billion words of the NOW corpus to assess very recent usage around the world.

Retrieving all relevant instances of the prepositional verb from the corpus is not a trivial task, since the tagger will have had problems distinguishing between verbal and nominal uses of the forms in question, and tag-based retrieval is thus unreliable. I therefore decided to search for the string "to leverage on" only, thus dramatically reducing recall, but at the same time retrieving (virtually) exclusively relevant verbal uses. The distribution over the 20 countries contained in the corpus is shown in Figure 3.

As is immediately obvious, the prepositional verb is almost non-existent in native-speaker varieties, with slightly higher frequencies in Australian and New Zealand English, albeit based on unreliably low absolute numbers. Similarly low frequencies are found in some ESL varieties (e.g. Pakistan and Bangladesh English), but most non-native varieties display markedly greater numbers. It thus seems possible to claim that the establishment of the innovative prepositional verb

Frequency by country (Return to frequency by year)

SECTION	FREQ	SIZE (M)	PER MIL	CLICK FOR CONTEXT (SEE ALL)
United States	6	585.8	0.01	
Canada	5	409.7	0.01	
Great Britain	6	431.1	0.01	
Ireland	0	147.3	0.00	
Australia	6	178.7	0.03	
New Zealand	2	57.3	0.03	
India	75	240.8	0.31	
Sri Lanka	4	11.5	0.35	
Pakistan	3	47.2	0.06	
Bangladesh	1	11.7	0.09	
Malaysia	162	33.9	4.78	
Singapore	143	45.8	3.12	
Philippines	39	67.1	0.58	
Hong Kong	21	7.9	2.65	
South Africa	23	165.4	0.14	
Nigeria	172	45.7	3.76	
Ghana	34	19.3	1.76	
Kenya	43	17.9	2.40	
Tanzania	1	6.4	0.16	
Jamaica	0	20.7	0.00	

Figure 3. The distribution of *to leverage on* over the 20 countries contained in the NOW corpus

to leverage on is a pan-ESL phenomenon. In addition, there appear to be some interesting areal preferences: Southeast Asian and (most) African countries display higher frequencies than South Asian countries.[17] Also, given the large number of first languages involved, the data displayed in Figure 3 further confirms the claim that substratum influence is an unlikely candidate for explaining the establishment of innovative prepositional verbs.

Naturally, I do not wish to claim that *to leverage on* developed independently in all of the ESL-varieties shown in Figure 3. We are clearly seeing the results of a globalized use of English in an interconnected world. However, the fact that native varieties do not seem to participate in this development is support for the claim that speakers of institutionalised second-language varieties are particularly open to such innovations. One fascinating avenue of research,

17. The status of South African English as native or ESL variety is difficult to assess (see e.g. Schneider 2007: 173ff.); it will therefore not be considered further in the present discussion.

which goes far beyond the scope of the present paper, is to investigate the type of areal preferences displayed in Figure 3. As a further case in point, consider Figure 4, which displays the distribution of the search for the prepositional verb *to mention about*.[18] This distribution is similar to, and at the same time also very different from, that of *to leverage on*. Again, native-speaker varieties exhibit low frequencies (indicated by the lighter shades of blue), and many of the hits are in fact false positives (e.g. *One point he mentions about Gomez caught my eye*). At the same time, the areal preference of the prepositional verb is radically different: *to mention about* is particularly frequent in (most) South and Southeast Asian countries, with African varieties exhibiting similarly low frequencies as native Englishes.

	CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
1	MENTIONED ABOUT	1338	166	44	200	37	55	27	163	81	116	67	99	102	60	26	17	18	17	7	24	12
2	MENTION ABOUT	970	140	52	131	23	59	20	118	72	75	42	45	77	23	36	7	10	4	14	11	11
3	MENTIONS ABOUT	161	26	2	24	1	4	1	25	16	8	13	6	13	5	4	4	2		4	1	2
4	MENTIONING ABOUT	153	9	3	18	2	4	2	36	13	20	10	8	15	4	5					1	3
	TOTAL	2622	341	101	373	63	122	50	342	182	219	132	158	207	92	71	28	30	21	25	37	28

Figure 4. The distribution of the prepositional verb *to mention about* over the 20 countries contained in the GloWbE corpus (search string: [mention] about)

Clearly, more work is needed to understand the processes that lead to such differences, and ideally, such research would involve both synchronic and diachronic data.

7. Concluding remarks and outlook

The present paper has focused on a number of prepositional verbs and their diachronic development in Singapore English. It is the first to do so on the basis of large amounts of real-time data. As a result, it has been possible to establish a link between quantitative data and theories of the evolution of post-colonial Englishes: the fact that *to request for* is used in almost a quarter of all possible instances of the construction by the year 1981 is strong evidence for a nativization process in action.

Not all of my results have been similarly clear. Thus, low frequencies of the literal use of *to enter into* make definite claims impossible, but the data neverthe-

18. This result has not been manually post-processed and therefore contains some false positives. However, the overall picture would not change as a result of such a manual clean-up.

less suggest that – in contrast to *to await for* – we are not dealing with a form that was retained from earlier uses of the input variety BrE. Finally, I focussed on *to leverage on*, which was determined to be a new, but quickly adopted, prepositional verb in SingE. This innovation also allowed me to open up the focus to more general considerations relating to the propensity of speakers of ESL varieties to opt for prepositional verbs as well as to intriguing areal preferences.

It has also become clear that there is still much that needs to be done. For example, we may ask whether the data used for the present study are in fact suitable for studying structural nativization. Newspaper language has been shown to be a particularly "agile" genre (cf. Hundt & Mair 1999), representing innovations in a language more quickly than other genres, but we are nevertheless dealing with an environment that is subject to heavy editing. In a country where exonormative standards are ferociously upheld by the government – cf. the Speak Good English Movement and its partial success in eradicating certain perceived non-standardisms – newspaper language may not be the best mirror of innovative forces at work. We therefore need access to a larger variety of data sources from different stages of the development of SingE, ideally including more speech-like genres and less heavily edited, semi-published works such as school yearbooks, school exams, etc. Furthermore, we need comparable diachronic corpora from different ESL varieties to generalise from variety-specific findings. Luckily, a number of projects are currently underway that will provide a sound basis for future studies. This, in turn, will give further momentum to the study of Englishes worldwide and enable researchers to link description and theory in even more convincing ways.

Acknowledgements

I would like to thank Gunnel Tottie for her helpful comments on a draft version of this paper.

References

Bamgbọṣe, Ayo. 1982. Standard Nigerian English: Issues of identification. In *The Other Tongue: English across Cultures*, Braj B. Kachru (ed), 99–109. Urbana IL: University of Illinois Press.

Bao, Zhiming. 2009. *One* in Singapore English. *Studies in Language* 33(2): 338–365. doi:10.1075/sl.33.2.05bao

Bao, Zhiming. 2010. A usage-based approach to substratum transfer: The case of four unproductive features in Singapore English. *Language* 86(4): 792–820.

Bao, Zhiming & Wee, Lionel. 1999. The passive in Singapore English. *World Englishes* 18(1): 1–11. doi:10.1111/1467-971X.00117

Biewer, Carolin, Bernaisch, Tobias, Berger, Mike & Heller, Benedikt. 2014. Compiling the diachronic corpus of Hong Kong English (DC-HKE): motivation, progress and challenges. Poster presentation at ICAME 35, Nottingham.

Bolton, Kingsley. 2002. *Hong Kong English Today. Autonomy and Creativity*. Hong Kong: Hong Kong University Press.

Borlongan, A. M. & Dita, S. N. 2015. Taking a look at expanded predicates in Philippine English across time. *Asian Englishes* 17(3): 1–8. doi:10.1080/13488678.2015.1069965

Brato, Thorsten. 2014. Compiling a historical written corpus of Ghanaian English: Methodological and theoretical considerations. Paper presented at the 20th Conference of the International Association for World Englishes, New Delhi.

Bruthiaux, Paul. 2010. The Speak Good English Movement: A web-user's perspective. In *English in Singapore: Modernity and Management*, Lisa Lim, Anne Pakir & Lionel Wee (eds), 91–108. Hong Kong: Hong Kong University Press. doi:10.5790/hongkong/9789888028436.003.0004

Davies, Mark & Fuchs, Robert. 2015. Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-based English Corpus (GloWbE). *English World-Wide* 36(1): 1–28. doi:10.1075/eww.36.1.01dav

Deterding, David. 2007. *Singapore English*. Edinburgh: EUP. doi:10.3366/edinburgh/9780748625444.001.0001

Fritz, Clemens. 2007. *From English in Australia to Australian English: 1788–1900*. Frankfurt: Peter Lang.

Garside, Robert. 1987. The CLAWS word-tagging system. In *The Computational Analysis of English: A Corpus-based Approach*, Robert Garside, Geoffrey Leech & Geoffrey Sampson (eds), 30–41. London: Longman.

Gonzales, Andrew. 1983. When does an error become a feature of Philippine English? In *Varieties of English in Southeast Asia*, Richard B. Noss (ed.), 150–172. Singapore: Regional Language Centre.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8: 41–60. doi:10.1075/eww.8.1.05gor

Gupta, Anthea Fraser. 1986. A standard for written Singapore English? *English World-Wide* 7(1): 75–99. doi:10.1075/eww.7.1.05gup

Gupta, Anthea Fraser. 1997. Colonisation, migration and functions of English. In *Englishes around the World: General Studies, British Isles, North America. Studies in Honour of Manfred Görlach*, Vol. 1 [Varieties of English around the World G18], Edgar W. Schneider (ed), 47–58. Amsterdam: John Benjamins. doi:10.1075/veaw.g18.08gup

Gupta, Anthea Fraser. 2008. The language ecology of Singapore. In *Encyclopedia of Language and Education*, Nancy H. Hornberger (ed), 2990–3002. Dordrecht: Kluwer. doi:10.1007/978-0-387-30424-3_224

Gupta, Anthea Fraser. 2010. Singapore Standard English revisited. In *English in Singapore: Modernity and Management*, Lisa Lim, Anne Pakir & Lionel Wee (eds), 57–89. Hong Kong: Hong Kong University Press. doi:10.5790/hongkong/9789888028436.003.0003

Hinrichs, Lars. 2006. *Codeswitching on the Web: English and Jamaican Creole in E-mail Communication* [Pragmatics & Beyond New Series 147]. Amsterdam: John Benjamins. doi:10.1075/pbns.147

Hundt, Marianne. 1998. *New Zealand English Grammar – Fact or Fiction? A Corpus-based Study in Morphosyntactic Variation* [Varieties of English around the World G23]. Amsterdam: John Benjamins. doi:10.1075/veaw.g23

Hundt, Marianne & Mair, Christian. 1999. 'Agile' and 'uptight' genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics* 4(2): 221–242. doi:10.1075/ijcl.4.2.02hun

Kachru, Braj B. 1982. Models of non-native Englishes. In *The Other Tongue: English Across Cultures*, Braj B. Kachru (ed.), 31–57. Oxford: Pergamon Press.

Kachru, Braj B. 1985. Standards, codification, and sociolinguistic realism: The English language in the outer circle. In *English in the World: Teaching and Learning the language and the literature*, Randolph Quirk & H. G. Widdowson (eds), 11–30. Cambridge: CUP.

Leech, Geoffrey, Hundt, Marianne, Mair, Christian & Smith, Nicholas. 2009. *Change in Contemporary English. A Grammatical Study*. Cambridge: CUP. doi:10.1017/CBO9780511642210

Leimgruber, Jakob R. E. 2013a. *Singapore English: Structure, Variation, and Usage*. Cambridge: CUP. doi:10.1017/CBO9781139225755

Leimgruber, Jakob R. E. 2013b. The trouble with World Englishes: Rethinking the concept of 'geographical varieties' of English. *English Today* 29(3): 3–7. doi:10.1017/S0266078413000242

Leitner, Gerhard. 1992. English as a pluricentric language. In *Pluricentric Languages: Differing Norms in Different Nations*, Michael Clyne (ed.), 179–237. Berlin: Mouton de Gruyter.

Merriam-Webster, online. <https://www.merriam-webster.com> (6 February 2017).

Moag, Rodney F. 1982. The life cycle of non-native Englishes: A case study. In *The Other Tongue: English Across Cultures*, Braj B. Kachru (ed), 270–288. Oxford: Pergamon Press.

Mukherjee, Joybrato & Gries, Stefan T. 2009. Collostructional nativisation in New Englishes: Verb-construction associations in the International Corpus of English. *English World-Wide* 30(1): 27–51. doi:10.1075/eww.30.1.03muk

Mukherjee, Joybrato & Hoffmann, Sebastian. 2006. Describing verb-complementational profiles of New Englishes: A pilot study of Indian English. *English World-Wide* 27: 147–173. doi:10.1075/eww.27.2.03muk

Mukherjee, Joybrato & Hundt, Marianne. 2011. *Exploring Second-language Varieties of English and Learner Englishes: Bridging a Paradigm Gap* [Studies in Corpus Linguistics 44]. Amsterdam: John Benjamins. doi:10.1075/scl.44

Nesselhauf, Nadia. 2009. Co-selection phenomena across New Englishes: Parallels (and differences) to foreign learner varieties. *English World-Wide* 30(1): 1–26. doi:10.1075/eww.30.1.02nes

Oxford English Dictionary, online. <http://www.oed.com> (6 February 2017).

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Schilk, Marco. 2011. *Structural Nativization in Indian English Lexicogrammar* [Studies in Corpus Linguistics 46]. Amsterdam: John Benjamins. doi:10.1075/scl.46

Schneider, Edgar W. 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79(2): 233–281. doi:10.1353/lan.2003.0136

Schneider, Edgar W. 2004. How to trace structural nativization: Particle verbs in World Englishes. *World Englishes* 23(2): 227–249. doi:10.1111/j.0883-2919.2004.00348.x

Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the World*. Cambridge: CUP. doi:10.1017/CBO9780511618901

Schneider, Edgar W. 2012. Exploring the interface between World Englishes and second language acquisition – and implications for English as a lingua franca. *Journal of English as a Lingua Franca* 1(1): 57–91. doi:10.1515/jelf-2012-0004

Schneider, Gerold & Gilquin, Gaëtanelle. 2016. Detecting innovations in a parsed corpus of learner English. *International Journal of Learner Corpus Research* 2(2): 177–204.

Schneider, Gerold & Zipp, Lena. 2013. Discovering new verb-preposition combinations in New Englishes. In *Corpus Linguistics and Variation in English: Focus on Non-Native Englishes* [Studies in Variation, Contacts and Change in English 13], Magnus Huber & Joybrato Mukherjee (eds). Helsinki: VARIENG. <http://www.helsinki.fi/varieng/series/volumes/13/schneider_zipp>

Schreier, Daniel. 2003. *Isolation and Language Change: Contemporary and Sociohistorical Evidence from Tristan da Cunha English*. Houndmills: Palgrave Macmillan. doi:10.1057/9780230505261

Schreier, Daniel. 2008. *St Helenian English: Origins, Evolution and Variation* [Varieties of English around the World G37]. Amsterdam: John Benjamins. doi:10.1075/veaw.g37

Sedlatschek, Andreas. 2009. *Contemporary Indian English. Variation and Change* [Varieties of English around the World G38]. Amsterdam: John Benjamins.

Singapore Department of Statistics. 2010. *Census of Population 2010 Statistical Release 1. Demographic Characteristics, Education, Language and Religion*. Singapore: Department of Statistics, Ministry of Trade & Industry, Republic of Singapore.

Stroud, Christopher & Wee, Lionel. 2010. Language policy and planning in Singaporean late modernity. In *English in Singapore: Modernity and Management*, Lisa Lim, Anne Pakir & Lionel Wee (eds), 181–204. Hong Kong: Hong Kong University Press. doi:10.5790/hongkong/9789888028436.003.0008

Tan, Siew Imm. 2016. Charting the endonormative stabilization of Singapore English. In *Communicating with Asia: The Future of English as a Global Language*, Gerhard Leitner, Azirah Hashim & Hans-Georg Wolf (eds), 69–84. Cambridge: CUP.

Tan, Siew Imm. 2013. Nativized prepositional verbs in Malaysian English from the perspective of language contact. *3L: The Southeast Asian Journal of English Language Studies* 3: 103–114.

Wee, Lionel. 2014. Linguistic chutzpah and the Speak Good Singlish Movement. *World Englishes* 33(1): 85–99. doi:10.1111/weng.12055

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Malden MA: Blackwell.

The development of infinitival complementation with or without language contact

Peter Slomanson
University of Tampere

Certain contact languages previously lacking a finiteness contrast have developed infinitival complementation. The late Old English and Sri Lankan Malay (SLM) constructions both involve *to*-infinitives seemingly based on prepositional phrases, specifically infinitival *to* + lexical verb in Old English and infinitival *nang* + lexical verb in SLM. There is no evidence, however, that these verbs were ever nominalized in SLM, and Los (2005) has argued that the apparently dativized forms we find in Old English belie the fact that their syntactic status was verbal and the constituents containing them clausal. The SLM infinitival prefix is etymologically irrealis, paralleling the subjunctives that the English *to*-infinitive progressively replaced. Identifying cross-linguistically parallel changes that are explainable based on textual attestations in one of the languages examined will aid in reconstructing the development of underattested languages that lack diachronic corpora.

Keywords: Sri Lankan Malay; finiteness; language contact; regrammaticalization

1. Introduction

The topic of this paper is the development of infinitival complementation in Sri Lankan Malay, a South Asian contact language. The comparative approach to be taken involves presenting a type of parallelism I have identified in that language with the development of the *to*-infinitive in non-finite complementation in English. English is a language whose diachrony is far better attested, and therefore potentially instructive for the study of languages whose diachrony is not similarly well-attested, lacking a significant or usable textual record. The diachronic data gap for Sri Lankan Malay is attributable to a history of rigid diglossia, effectively excluding the changing spoken language from the available written record, so that diachronic research has largely been restricted to plausible theorizing. That is a common logistical dilemma in contact language research, but the research itself is

DOI 10.1075/slcs.195.11slo

nevertheless potentially illuminating, as it provides us with a cross-linguistically broader range of parallels on the ways in which sentential complementation patterns change over time.

Based on the data gap between unwritten contact varieties and better-attested languages such as English, variation and change trajectories that are clearly attested or straightforwardly reconstructable for languages such as English are, logically, indispensable. The contact language data also contribute to this area of inquiry, however, providing greater scope for our knowledge of universal processes, including those in which collective second language acquisition is an important factor influencing grammatical outcomes. The more abundant data are needed to facilitate empirically-informed modeling of the grammatical history of those languages that lack large diachronic corpora, but that may have phenomena in common with English and other well-documented languages.

2. Cross-linguistic variation in interclausal symmetry

If we examine a random cross-section of languages, we will notice that complement clauses and non-root clauses in general come in a range of types. One broad cross-linguistic contrast is between languages of what I will call type A and languages of what I will call type B. Type A languages, such as canonical pidgins and creoles, and varieties of Chinese and Malay, among others, that are mostly isolating or at least highly analytic, demonstrate a kind of interclausal symmetry, concatenating clauses paratactically, or with the use of conjunctions that are not subordinating conjunctions, because they do not modify anything in the syntactic or morphological content of the clause they introduce.

Type B languages, on the other hand, such as English, Irish, Finnish, Tamil, and so on, feature complements that are clearly distinct from root clauses in form, one option being non-finite complements, including infinitives and nominalizations, another option being finite complement clauses, such as Germanic *that*-clauses, that differ syntactically from their matrix clauses.[1] [2] The infinitival complement

1. This intersentential contrast is typically found in such alternations as a contrast in the position of the finite verb or in the syntactic relationship of the finite verb to negation and sentence adverbs.

2. The criteria I am using here are morphosyntactic, that is essentially inflectional criteria. From a functional perspective as in Stassen (1985), type A contains only balanced clauses, meaning that complement clauses could, due to their form, function as independent clauses. The B languages may feature complement clauses containing verbs that are balanced, deranked, or both. Deranked verb forms are in some sense morphologically underspecified,

clauses in the B languages (the non-symmetrical languages) may contrast still more obviously with what we find in finite root clauses, the infinitival verb not being marked in any way for tense, nor usually for subject-verb agreement.[3] In this non-symmetrical type of language, there are different potential strategies for marking the non-finite status of the affected non-matrix verb, for example, with an ordinary preposition or a syntactically reanalyzed preposition such as *to* as an infinitive marker, with an infinitival suffix, or with a combination of pre- and post-verbal markers. These contrasts with verbs in root contexts are often morphological in this way, and not necessarily visible from word order.

3. The historical English analogy with Sri Lankan Malay

The interaction between different potential formal complementation strategies is interesting from early on in the history of English, since there was a time segment, lasting into the Middle English period, during which subjunctive *that*-clauses across the lexicon were progressively replaced by *to*-infinitives. The standard claim is that it is the older infinitive type with which the *to*-infinitive was in competition. The *to*-prefix would indeed replace the older infinitive, eventually becoming a free-standing infinitive marker, the one still used in present-day English.[4] Los (2005), in addition to calling into question the centrality of competition between the original suffixal infinitives (i.e. with *-an*) and the rising *to*-infinitive and the eventual loss of the old infinitive type in the Middle English period, also challenges the claim that *to*-infinitives were nominal, so it is inaccurate to think of them as dativized verbs. She also claims that their dative suffixes were fossilizations in the relevant period. It would be the development of a parallel construction, in the form of a seemingly prepositional (postpositional) phrase containing an apparent dative marker that Sri Lankan Malay behaved in its diachronic development the way English did.

---

so that they cannot function as independent or matrix verbs, according to the morphosyntactic criteria of the language such as finiteness, but this functional distinction is not cross-linguistically limited to the finite/non-finiteness contrast.

3. European Portuguese has infinitives that agree with their subjects.

4. The justification for not treating infinitival *to* as a prefix on lexical verbs is the fact that adverbs can separate particle *to* from the non-finite VP that is its syntactic complement. In other distributional evidence, when a matrix verb subcategorizes for a *to*-infinitive, the lexical segment in the infinitival construction can be deleted under identity.

Based on the evidence currently available to us, there is in fact no more reason to analyze Sri Lankan Malay infinitives as nominalizations, nor to treat infinitival complements in Sri Lankan Malay in that sense as constituents headed by adpositions than does Los in her analysis of infinitival complements in Old English.[5] Part of what shows us this is the pre-verbal marker on Sri Lankan Malay infinitives, an apparent inflection. Etymologically, it is a phonologically-reduced volitive modal in Malay contact varieties spoken in Indonesia, the lexical source varieties for Sri Lankan Malay (see, for example, van Minde 1997).[6] English has never had verbs inflected for modality; however, the semantics of the free-standing infinitival marker in English and the bound infinitival marker in SLM both suggest unrealized events. In SLM, this is simply that much more visible, given the apparent etymology of the functional element. SLM has both the subjunctive-like properties attributed by Los to English *to*, in addition to the adposition, the postposition *nang*. This *nang* is also, like *to*, a dative/allative marker; however, it is not clear that in association with the lexical verb, it has any semantic or syntactic features that would suggest that the infinitival verb is a nominalization.

4. The view that restoration of infinitival marking is improbable

There is a body of language contact research in which an assumption is made that, whatever the external scenario yielding a particular contact language variety, a type A language, that is the type featuring cross-clausal symmetry, a common creole type, is unlikely ever to become a type B language, whereas the converse is certainly not true. The extent to which the A/B contrast applies to any variety of

5. Slomanson (2006) analyzed the Sri Lankan Malay as a nominalization, due to their apparent dative clitic (the analog of English to) and the fact that certain infinitival complements instead receive accusative case-marking, as in Tamil. In subsequent work, I retracted this position, based in part on the fact that the accusative-marked construction is vanishingly rare, and may have constituted a spontaneous transfer from Tamil/Shonam.

6. I am using the phonologically-reduced (from *mau* to *mo* to *mə*) future/irrealis marker in Ambonese Malay as an example of a tendency in Malay contact varieties, without claiming that Ambonese Malay was the specific source, although eastern Indonesian Malay contact varieties generally are likely to have been. Smith and Paauw (2006: 173) treat Sri Lankan Malay *mə-* as being derived from the active transitivizing prefix in other Malay varieties, which is homophonous with Sri Lankan Malay *məN-*. Semantically, this is not well-motivated, particularly in comparison with the Ambonese Malay analogy. Nordhoff (2009: 280) points out that "the development from a transitivizing morpheme to an infinitive is not yet attested in the literature on grammaticalization paths".

infinitival complementation in contact languages is contentious, as there is a type of uninflected verbal construction that is common in creoles and other otherwise symmetrical languages (in the sense discussed here) that has frequently been characterized as infinitival in the creole linguistics literature. This characterization [of those creole complements as infinitives] has been challenged in Bickerton (1984) and Mufwene (1989). In a strict (for example Bickertonian) view of what radical language contact entails, if a language loses its morphological infinitives, or if it perhaps never had them, then it will be difficult and unlikely to get new ones, so to speak.[7] Mufwene emphasizes a morphological basis for infinitival status, a criterion that Sri Lankan Malay, in its expanded functional and inflectional complexity, happens to fulfill. His challenge is to the view that non-inflectional infinitives can and do develop in creoles, usually introduced by a reflex of what would be *for* in their respective lexical source languages. We can find a relatively early, though less categorical precursor to the Bickertonian view that infinitival morphology is unlikely to be restored from earlier literature on bilingual language contact, by revisiting a suggestion in Weinreich (1953) that free-standing elements are much easier to borrow than bound morphology. In reality, not just infinitives generally, but bound infinitival morphology can be regrammaticalized, and this is not limited to purposive infinitives, whose semantic function is unambiguous. This observation is based on evidence from a relatively small subset of contact languages, sometimes called converted languages, of which Sri Lankan Malay is one, in which the conversion from one language type to another is largely morphosyntactic.[8] Where regrammaticalization does take place, the actual phonological shapes themselves need not be borrowed, and often they are not. But the development of a morphological infinitive is still possible using lexical material that is not borrowed, just reassigned. What we find in that case is a new language with a lexical inventory from one language family and the morphosyntax of another.

---

7. A claim that infinitives are by their nature highly marked and for that reason unlikely to be restored after loss (due for example to pidginization of an earlier grammar that contained them) is difficult to reconcile with evidence on child first language acquisition. That evidence has shown that in languages with both finite and infinitival verbs, children pass through a stage in which finite verbs are not used, but infinitival verbs in their appropriate syntactic context with respect to verb-complement ordering, are used (see, for example, Blom 2003).

8. I am treating English here as a non-contact language, but contact language status and non-contact language status are really graded rather than discrete. Perhaps conventional languages such as English, historically subject to contact effects, including substratal ones, versus radical contact languages, such as Sri Lankan Malay, would be a more precise description of the contrast.

The discussion in the papers by Bickerton and Mufwene pertains to canonical creoles, although the broader generalization has often been taken for granted in contact linguistics generally. It may be taken to follow from assumptions about the nature of linguistic complexity, as in Weinreich's suggestion that inflectional morphology can be difficult to model based on the grammar of another language, and the claim that the morphology and syntax involved in infinitival complementation and other clausal asymmetries should be treated as a type of complexity that will attrite under contact conditions and not return. However, it often does attrite, although not necessarily. The reason for the tendency to attrite is that the relevant morphology is easily filtered as a kind of grammatical "noise" by second language acquirers under constrained naturalistic acquisition pressure, for example pidginization contexts, but also in various immigration and other scenarios. A related assumption is that the asymmetries themselves have relatively little communicative salience. But an assumption that finiteness asymmetries, including but not limited to infinitival complementation, have minimal or non-existent pragmatic functions is not cross-linguistically valid. Infinitives and participles may have clear discourse functions, including the marking of aspectual information, such as event sequencing, simultaneity, and the irrealis status of events and states. Any communicative-pragmatic value in the finite/non-finite contrast does have some bearing on whether or not the contrast will be regrammaticalized at a later stage, after having been lost at an earlier one.[9]

5. The development of infinitival complementation in Sri Lankan Malay

With respect to Sri Lankan Malay, how and why did infinitival morphology develop in a language that once had an essentially flat relationship between clauses in complex sentences? This change does not follow from contact with English, but rather from contact with indigenous South Asian languages, and particularly

9. Slomanson (2016) discusses outcomes of bilingual language contact scenarios, specifically whether the development of infinitival complementation and participles or other non-finite adjuncts can be attributed at least in part to the discourse culture in which the contact language develops. If appropriate conditions are met, then changes (or the absence of changes) can potentially be shown to depend on communicative-pragmatic factors, since such factors are the primary target for ordinary bilingual speakers in a way that morphosyntactic features are not. Sri Lankan Malay, as one example, has shifted in its short history from type A to type B, in spite of the standard assumptions. Other Malay contact varieties did not and do not have a finite/non-finite contrast, let alone infinitival complementation. Sri Lankan Malay now has an obligatory one, involving explicit contrastive tense marking and infinitival morphology.

with Tamil as spoken, highly non-standardly, by Sri Lankan Muslims, in a variety traditionally referred to by them as Shonam.[10] This does not mean that the new Malay grammar perfectly replicates the grammar of Shonam, nor does the new grammar necessarily align in other ways with the grammar of English, historical or present-day.

This new Malay syntax is not close to the syntax of English across the board. The parallel is only with respect to infinitival complements and the way the relevant construction developed, in English and in Sri Lankan Malay. In infinitival complementation, an important *difference* is the fact that Sri Lankan Malay did not develop Exceptional Case Marking constructions of the type in (1), in which a non-nominative object is the subject of the embedded complement clause, but it does have subject control constructions of the type in (2).

(1) Miflal wanted [me to cook the rice]

(2) Miflal tried [to cook the rice]

In fact, the Sri Lankan Malay counterpart of *try* cannot ever take a finite complement any more than English *try* can, as in the ungrammatical example in (3).

(3) *Miflal tries [that he cooks the rice]

These facts about the syntax of verbs of this type and their complements likely have to do with the irrealis or forward-looking semantics of the infinitival complement, as selected by the matrix verb *try*.

Examples (4) through (6) illustrate the radical changes to Malay complementation in Sri Lanka, as well as the parallel with English. Looking at the parallel examples with identical meaning in (4) through (6), with the complement clauses in square brackets, we see that although (4) and (5) are, not surprisingly, lexically similar, (5) and (6) are highly similar grammatically. Of course the adjacency requirement in Shonam highlights a contrast in the relationship of the modal to its infinitival complement; however, there is also an affixal variant of each modal, that can occur as a bound atonic clitic on the left of any verb such as *pi* in (5).

---

10. The traditional written version of this Tamil variety has been referred to as Arabu-Tamil, as it followed the historical tendency of Muslim ethnic groups to write the languages they use for in-group functions in modified Arabic script. The tendency in the last century was to abandon that script and to produce texts in Tamil orthography. The tendency in Malay, also a written language in Sri Lanka, was to abandon the written word in the community's language entirely. It should be noted that there were many bilingual (Shonam-Malay) texts produced in the nineteenth century in modified Arabic script, and that the first Malay-language newspaper in the world was published in Sri Lanka, in modified Arabic script.

The near parallel is clear, as is the Sri Lankan (Shonam) model for the Sri Lankan Malay construction.[11]

Malay

(4) ***Miflal mau*** [***pi ke ruma***] ***sekarang.***
Miflal want go to house now
'Miflal wants to go home now.'

Sri Lankan Malay

(5) ***Miflal*** [***ruma-nang mə-pi nang***] (***karang***) ***kəmauan.***
Miflal house-to INF-go to (now) want
'Miflal wants to go home now.'

Shonam

(6) ***Miflal-ukku*** [***viʧ-ukku poo-k-***(****ippo***) ***oonum***].
Miflal-DAT house-to go-INF-(*now) want
'Miflal wants to go home.'

The speaker of Sri Lankan Malay can and does select stylistically between the pre-verbally bound modal option (7) and the option in which the modal is a free-standing predicate (8).

Sri Lankan Malay

(7) ***Miflal ruma-nang bole-pi.***
Miflal house-to MOD-go
'Miflal can go home.'

Sri Lankan Malay

(8) ***Miflal*** [***ruma-nang mə-pi nang***] ***si-bole.***
Miflal house-to INF-go to PST-MOD
'Miflal can go home.'

In Shonam, morphological infinitives are embedded within agglutinative constructions such as (6) that require the post-nominal adjacency of the modal and the infinitive. In the Sri Lankan Malay construction involving an affixal modal,

---

11. The example in (5) happens to resemble the Sinhala construction involving a free-standing predicate modal with an infinitival complement to its left.

Sinhala

(i) ***Miflal*** [***gedərə-tə yannə***] (***dæng***) ***oonə.***
Miflal house-to go-INF (now) want
'Miflal wants to go home now.'

In Slomanson (2011), linguistic arguments are presented for treating Shonam as having served as the primary model for infinitival complementation and for other developments in the morphosyntax of Sri Lankan Malay, with Sinhala influence as adstratal or secondary. Slomanson (2013) provides sociohistorical arguments for this primacy.

the modal is bound to the left edge of a finite lexical verb; however, the finite and non-finite clauses are separable in the modal predicate construction. For a generative analysis of this construction, see Slomanson (2008). In (5), the grammatical modeling of Sri Lankan Malay on Shonam can be seen from this evidence. These examples show us that Sri Lankan Malay is now left-branching. They also show us that the language has developed bound functional morphology of various types, and that a clausal complement is now most likely to appear to the left of a matrix predicate, just as nominal object complements precede verbs in pragmatically-unmarked contexts within clauses (OV order), and complements of adpositions always precede their heads (obligatorily postpositional PPs). This makes Sri Lankan Malay appear to be a canonical SOV language, with all major constituents head-final, just as they are in Shonam, although unlike in Shonam, bound functional morphology associated with tense and mood contrasts precedes the verb.[12] The subject may appear to the left of this biclausal construction with an infinitival complement (as with the subject *Miflal* in 5). However, it is typically the first constituent in a sentence lacking adjunct clauses. The infinitival clause is center-embedded in (5), which is the unmarked order.[13]

---

12. Tense morphology was not present in (5) because *kəmauan*, meaning 'want', happens to be a nominal predicate that takes an infinitival complement. In (ii) we see an open class verbal predicate, *liyat*, meaning *try*, which takes an obligatory tense prefix.

Sri Lankan Malay

(ii) ***Miflal*** [***ruma-nang*** ***mə-pi nang***] (***karang***) ***arə-liyat.***
Miflal house-to INF-go to (now) PRS-try
'Miflal wants to go home now.'

13. One of the contact linguistic innovations in Sri Lankan Malay involving non-finite verbs is the introduction of a new Malay version of the South Asian conjunctive participle construction in adjunct clauses, standardly translated as "Having done x, …". It is the significance of clause dislocation in order to focus temporally non-primary events that led to this innovation. If a speaker dislocates a participial adjunct clause denoting an event to the right edge of the sentence in order to focus the dislocated event as new information, then the non-primary temporal status of the dislocated clause in a sequence of events remains marked. The focusing of a constituent by dislocating it to the right edge of a clause, typically the nominal argument of a verb, is replicated in the dislocation of a participial adjunct to the right edge of a sentence. Hence the function of morphologically marking non-finite status, since matrix verb tense morphology has the function of demonstrating most recent (i.e. primary) event status. This facilitates the interaction of differential temporal status (primary and non-primary) and focal status (new information versus old or given information).

Sri Lankan Malay

(iii) ***iskuul*** ***na*** ***a(bi)s-pi*** ***mulbar*** ***a(bi)s-blajar,*** ***Miflal*** ***attu=nyanyi*** ***su-tulis.***
school P ASP.NFN-go Tamil ASP.NFN-learn Miflal IND=song PST-write
'Having gone to school, (and then) having learned Tamil, Miflal wrote a song (it).'

Notice that the Sri Lankan Malay sentence in (5) is unambiguously biclausal in an English-like way and features infinitival morphology on the lexical verb in the complement clause. The post-verbal infinitive marker is *nang*. This *nang* happens to resemble infinitival *to* in English, in its function and its distribution. The apparent source of Sri Lankan Malay *nang* as an infinitive marker is the homophonous Javanese allative adposition (Slomanson 2006, 2011), which is analogous with the English preposition *to*. This is reminiscent of that stage in the history of English at which verbs in infinitival constructions appeared to be datives. Los (2005) takes the clausal status of those constructions to have developed early and rejects the idea that infinitival *to* is a reanalyzed preposition. The frequency of *to*-infinitives eventually increased dramatically, with the diffusion of new *to*-infinitives through (though not throughout) the English lexicon, replacing subjunctive *that*-clauses, and yielding the contrast that we can see in present-day English in (10) through (13).

The core idea then is that *to*-infinitive constructions were not prepositional phrases, since the syntax of the ostensibly dativized lexical verb was not nominal, and the verbal complements of Old English *to* were not actually nominalizations. Several types of evidence discussed in Los apply equally to the Sri Lankan Malay infinitive. The most robust of these is the fact that nominal complements of infinitives in both languages are regularly marked with accusative case (9), rather than, for example, genitive case.

Sri Lankan Malay

(9) ***Miflal nasi=yang mə-makan nang kəmauan.***
Miflal rice=ACC INF-eat to want
'Miflal wants to eat the rice.'

Instantiating such a claim is straightforward for Sri Lankan Malay, because case has been morphologized in that language, whereas these were completely unmarked in the Malay lexical source language. It is also not possible to modify infinitives in any of the ways that one would modify an actual nominal constituent, either adjectivally, with quantifiers, or with (in)definiteness marking. In this sense, for Sri Lankan Malay, the *only* evidence for nominal status is the etymology of the infinitival marker, and the fact that a homophonous element continues to mark the dative/allative status of genuinely nominal constituents.[14]

Sri Lankan Malay

(iv) ***iskuul na a(bi)s-pi Miflal attu=nyanyi su-tulis, mulbar abis-blajar.***
school P ASP.NFN-go Miflal IND=song PST-write Tamil ASP.NFN-learn
'Having gone to school, Miflal wrote a song, having learned Tamil.'

14. In spite of the usefulness and legitimacy of this type of argument, the nominal or verbal status of complement clauses cannot be treated as a binary alternation, since it is

In the following present-day English examples, we see the contrast between semantically-related sentences alternating between an object control *to*-infinitive complement and a subjunctive complement, the selection being a property of the respective lexical verbs.

(10) Miflal told Rishan/him to go.

(11) *Miflal demanded Rishan/him to go.

(12) *Miflal told that Rishan/he go.

(13) Miflal demanded that Rishan/he go.

The verb *tell* is amenable to infinitival complementation, as seen in (10). The abstract subject in its infinitival complement is controlled by the matrix object (i.e. of *tell*). By contrast, the verb *demand*, perhaps due to its non-Germanic etymology, persists in selecting (requiring) a mandative (subjunctive) complement, rather than an infinitival one. In Sri Lankan Malay, monovalent control sentences (i.e. with a single matrix argument and subject control) are grammatical (for example with counterparts of *try*). The type of bivalent sentence often used to illustrate the

---

cross-linguistically graded in practice. One language whose grammar makes this quite clear is Irish, since that language offers both a type of complement that clearly bears the features of an NP, including nominal complements as genitives, and one that is straightforwardly clausal, though the traditional grammatical descriptions describe both types as containing a so-called verbal noun.

(v) ***Tá*** ***na*** ***páistí*** ***ag*** ***labhairt*** ***na*** ***Maláisise***
AUX.PRS DET.PLU children at speak DET.GEN Malay.GEN
sa ***tsráidbhaile.***
in-the village
'The children are speaking Malay in the village.' (lit. at the speaking of the Malay)

(vi) ***Tá*** ***ar*** ***na*** ***páistí*** ***an*** ***Mhaláisis*** ***a*** ***labhairt***
AUX.PRS on DET.PLU children DET.PLU Malay INF speak
sa ***tsráidbhaile.***
in-the village.
'The children must speak Malay in the village.'

The construction in (v) features the Celtic periphrastic progressive construction that is comparable to the periphrastic progressive construction in English. The verbal complement assigns visible genitive case to its nominal complement, *na Maláisise*, demonstrating in that way, and in the linear constituent order, that the clause-like constituent *ag labhairt na Maláisise* is a prepositional phrase. The complement clause in (vi), with its object-verb order and (abstract) accusative case assigned to the nominal complement of *labhairt*, is straightforwardly verbal (i.e. infinitival).

contrast between subject and object control (for example, *promise* and *persuade*, respectively) does not take infinitival complements at all. The generalization is that directive and commissive verbs describe speech acts, with the directed and committed actions appearing in their respective complements. In the [Sri Lankan] discourse culture, the action that the matrix verb introduces as its complement is actually quoted.[15] The quotative marker, which is clause-final, in keeping with the unmarked left-branching (complement-head) order of constituents in Sri Lankan languages, is the only available complementizer for finite complements. Matrix verbs that describe speech acts are therefore compelled, as a function of their semantic features, rather than of subcategorization features irrespective of lexical semantics, to take finite complements that may nevertheless encode the irrealis sense of English infinitives. That irrealis meaning correspondingly may be the reason that only infinitival complements are grammatical with etymologically Germanic matrix verbs of this type.

We can elaborate on the relationship of the *nang*-infinitive in SLM to infinitival complements in the Shonam model language. If we compare (5) and (6) again,

Malay

(4) ***Miflal mau*** [***pi ke ruma***] ***sekarang.***
Miflal want go to house now
'Miflal wants to go home now.'

Sri Lankan Malay

(5) ***Miflal*** [***ruma-nang mə-pi nang***] (***karang***) ***kəmauan.***
Miflal house-to INF-go to (now) want
'Miflal wants to go home now.'

Shonam

(6) ***Miflal-ukku*** [***viɬɬ-ukku poo-k-(*ippo) oonum***].
Miflal-DAT house-to go-INF-(*now) want
'Miflal wants to go home.'

we will see that the word for 'house' in (6) bears the case suffix *-ukku*. This is a nominal affix used in dative and allative contexts.[16] The Shonam infinitive is

---

15. Sri Lankan Malay

(vii) ***Go e-susa (ada) lompe fon e-ilang (ada) kata.***
1S ASP-worried (AUX) your phone ASP-lose (AUX) QUT
'I was worried that your phone was lost.'

16. However, it can also attach to gerundial lexical verbs that have purposive interpretation. These are not actual complements, but adjuncts. However, the analogous structure in Sri Lankan Malay has been extended to infinitival complements.

marked with a distinct infinitival suffix, however, and not with *-ukku*, whereas in SLM, we see *nang* the allative postposition cum dative case clitic, and *nang* the apparent analog of infinitival *to* in English, which in spite of their shared etymology, are syntactically differentiated. We can see both types of *nang* in (5). In Sri Lankan Malay, the functional semantics of the adposition that became a dative/allative case clitic was once extended to purposive adjuncts, which is actually completely in line with the Shonam pattern.[17] Complexifying contact languages are susceptible to generalizing structural patterns as a means of extending available resources. In that sense, Sri Lankan Malay has generalized the verb + *nang* construction to straightforward infinitival complements. In Shonam, there are actually different types of infinitival adjunct, with dative-marked verbs in adjunct clauses most likely to be purposive, while infinitival complements bear specifically infinitival morphology. In Sri Lankan Malay, moreover, the construction with *nang*, the analog of English infinitival *to* is actually the only option for infinitival complementation, whereas Shonam has more than one type of infinitival suffix. The sequence of events was surely triggered by contact, given the interclausal symmetry of the original colloquial Malay varieties. However, the generalization appears to follow a trajectory that has been treated as universal in typological literature, for example Haspelmath (1989), in which it is the initial allative origin of infinitival *to*, leading to a purposive function, that the author treats as universal. That is particularly clear in a converted language such as SLM, since the function of *nang* in the Javanese source language is not syntactically ambiguous.

Another observation about (5) is that the infinitival marking is actually discontinuous. The post-verbal element *-nang* is there, but but we also find the pre-verbal prefix, *mə-*. These two infinitival markers together support the features encoded by English *to*, including the irrealis or forward-looking meaning associated with subjunctives. There is also the morphosyntactic fact that *mə-* and modal prefixes are in complementary distribution, whereas free-standing modals can be rendered as infinitives, as was possible in English before modals were reanalyzed as INFL elements at a later stage in the history of English. One way of looking at the developments in the history of English is that infinitival *to* came to be associated with a structural position that epistemic, deontic, and temporal properties of verbs are associated with, ultimately as if *to* itself were a

17. Shonam

(viii) ***ange saapiDr-ad-ukku poor-een.***
there eat-VBN-DAT go-1S
'I'm going there (in order) to eat.'

modal, given the complementary distribution that obtains between modals and *to* in present-day English.

There is evidence for a similar development in Sri Lankan Malay, as evidenced by the complementary distribution. The resulting grammatical options, however, are somewhat different. English has bare infinitives in clauses headed by modals, whereas the Sri Lankan Malay strategy for deploying modals with this type of predication is to create and favor a biclausal structure with genuine *nang*-infinitives, in which the infinitival clause is selected by a matrix modal predicate which is negated independently with the finite negation marker *tər-*. Free-standing modals have the same etymology as the modal prefixes, but their syntactic status is completely different. English, in contrast with Sri Lankan Malay, no longer has modals that can occur as infinitival complements (as in **to can*), consequently circumlocutions such as "to be able to" must be used instead. For Sri Lankan Malay, we can see the two semantically equivalent options in (14) and (15). The sentence in (14) is monoclausal, with the subject taking an IP complement. The sentence in (15) is biclausal, with the matrix predicate modal taking an infinitival clause as its complement.

Sri Lankan Malay

(14) *Miflal* [*nyanyi attu bər-bilang*].
Miflal song IND MOD-say
'Miflal can sing a song.'

Sri Lankan Malay

(15) *Miflal* [*nyanyi attu mə-bilang nang*] *bole*.
Miflal song IND INF-say to MOD
'Miflal can sing a song.'

Sri Lankan Malay *mə-*, the infinitive marker, is derived from an irrealis marker in other contact Malay varieties that is itself derived from a volitive marker, *to want* (Slomanson 2006, 2011). The infinitive marker *mə-* formally conveys the irrealis sense that we associate with English infinitival *to*, whereas *nang*, a dative/allative case marker on noun phrases, reflects the fact that the infinitival complement construction is the generalization of a purposive adjunct construction. In that sense, *nang* may be a vestige of an older complementation strategy, consequently we might expect *nang* to be more frequently deleted than *mə-* is. This is in fact the case. This ultimately continues to resemble the trajectory that historical English went through.

A summary of the functional comparison of the relevant items follows. I have grouped *to* with *mə-,* based on the historical analysis of *to*-infinitives in Los (2005), in which *to* assumes the syntactic and functional-semantic status of a modal.

-nang: The Sri Lankan Malay suffix *nang* is a dative/allative clitic in noun phrases. However, the *nang* we find associated with verb phrases is an infinitival marker, the result of reanalysis without replacement of the original dative marker in noun phrases. This is variably deleted.
-ne: The Old English suffix, etymologically a dative marker in nominal contexts, was variably deleted from infinitives.
mə-: The Sri Lankan Malay infinitival prefix *mə-* is semantically irrealis and cannot occur in the same clause as a bound modal morpheme, all of which have tense features (Slomanson 2008). Instead the infinitival clause may occur as the complement of a matrix predicate modal.
to: The English infinitive marker *to* was and is semantically irrealis or forward-looking (depending on the semantic properties of the associated verb) and cannot occur in the same clause as a modal. Verbs taking this type of infinitival complement gradually replaced verbs taking semantically similar subjunctive complements.

6. Conclusion

Sri Lankan Malay, a 300–350 year old contact language, has developed a *to*-infinitive in which the Javanese allative preposition *nang* has been regrammaticalized as an infinitive marker, at the same time as it has become a (post-nominal) dative marker, with the two homophonous morphemes co-existing in the language. In this sense, the development is strongly reminiscent of the development of the *to*-infinitive in Old English, frequently analyzed as having been based on prepositional phrases in which the lexical verb is a dativized nominalization. The Sri Lankan Malay facts align better with the alternative view presented in Los (2005) that what appears to be a prepositional phrase is in fact simply a clause, due to the etymology and semantics of the *mə-* marker, a morphologically-reduced variant of an irrealis marker in the original Malay lexical source language. In Old English, the rise of the *to*-infinitive corresponds with the decline of an existing subjunctive construction. The etymology of *mə-* in Sri Lankan Malay supports the semantics of a subjunctive construction, and while purposive meanings are possible, there is no corpus evidence that would show us conclusively whether purposive meanings preceded the generalized infinitival complementation function that the [PRO *mə-*V-*nang*] construction now serves. The fact that English-type control sentences are always grammatical, while Exceptional Case Marking constructions are not, also parallels, the Old English facts.

Diachronic comparisons of well-attested and weakly-attested languages are useful in contact linguistic research, since it is from well-attested trajectories that we learn what types of changes are likely to take place under which internal and

external conditions. Without that knowledge, our tools for hypothesis building on the diachrony of weakly-attested contact languages would be far more limited. In that sense, explicit comparisons of well-attested and weakly attested language that have undergone comparable developments such as the instantiation of new forms of infinitival complementation will continue to be useful in linguistic research. That is what I have attempted to demonstrate in this preliminary comparison of infinitival complementation patterns and their development in two unrelated languages.

Abbreviations

1S	first person singular
AUX	auxiliary
ASP	aspect
DAT	dative
INF	infinitive
MOD	modality
NFN	non-finite
PRS	present tense
QUT	quotative
VBN	verbal noun

References

Bickerton, Derek. 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7: 173–221. doi:10.1017/S0140525X00044149

Blom, Wilhelmina Bernardina Theodora. 2003. *From Root Infinitive to Finite Sentence: The Acquisition of Verbal Inflections and Auxiliaries*. PhD dissertation, University of Utrecht.

Haspelmath, Martin. 1989. From purposive to infinitive: A universal path of grammaticization. *Folia Linguistica Historica* 10(1–2): 287–310.

Los, Bettelou. 2005. *The Rise of the To-Infinitive*. Oxford: OUP. doi:10.1093/acprof:oso/9780199274765.001.0001

Mufwene, Salikoko. 1989. On the so-called 'infinitive' in Atlantic creoles. *Lingua* 77: 297–330. doi:10.1016/0024-3841(89)90043-0

Nordhoff, Sebastian. 2009. A Grammar of Upcountry Sri Lanka Malay. PhD dissertation, University of Amsterdam.

Slomanson, Peter. 2006. Sri Lankan Grammars: Lankan or Malay? In *Structure and Variation in Contact Languages* [Creole Language Library 29], Ana Deumert & Stéphanie Durrleman (eds), 135–158. Amsterdam: John Benjamins. doi:10.1075/cll.29.08slo

Slomanson, Peter. 2008. The perfect construction and complexity drift in Sri Lankan Malay. *Lingua* 118(10), 1640–1655. doi:10.1016/j.lingua.2007.08.009

Slomanson, Peter. 2011. Dravidian features in the Sri Lankan Malay verb. In *Creoles, their Substrates, and Language Typology* [Typological Studies in Language 95], Claire Lefebvre (ed.), 383–409. Amsterdam: John Benjamins. doi:10.1075/tsl.95.22slo

Slomanson, Peter. 2013. Known, inferable, and discoverable in Sri Lankan Malay research. In *The Genesis of Sri Lanka Malay - A Case of Extreme Language Contact*, Sebastian Nordhoff (ed.), 85–119. Leiden: Brill.

Slomanson, Peter. 2016. New information structuring processes and morphosyntactic change. In *Information Structure and Spoken Language in Cross-Linguistic Perspective*, M. M. Jocelyn Fernandez-Vest & Robert D. Van Valin Jr. (eds), 305–326. Berlin: Mouton de Gruyter.

Smith, Ian & Paauw, Scott. 2006. Sri Lanka Malay: creole or convert? In *Structure and Variation in Contact Languages* [Creole Language Library 29], Ana Deumert & Stéphanie Durrleman (eds), 159–182. John Benjamins. doi:10.1075/cll.29.09smi

Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Blackwell.

van Minde, Don. 1997. *Malayu Ambong: Phonology, Morphology, Syntax*. Leiden: Research School CNWS.

Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York NY: Linguistic Circle of New York. (reprinted in 1979 by Walter de Gruyter).

Anglicising Finnish complementation? Examining the *rakastan puhua* ('I love to speak') structure in present-day Finnish

Pertti Hietaranta
University of Helsinki

The paper addresses the question as to whether the complementation system of Finnish is currently undergoing a change which will ultimately lead to Finnish syntax recognising *rakastan puhua* ('I love to talk/speak') as a valid instantiation of the syntactic structure in question in addition to the traditional construction *rakastan puhumista* ('I love talking/speaking'). At present, the former type is distinctly a marked type of expression whereas the latter variant is equally unmarked. The paper argues that while the emergence of the new type in Finnish may be attributable to the influence of the corresponding infinitival construction in English via translation, this is so only to some extent, and that the change Finnish syntax is undergoing in this respect is mainly due to some sociolinguistic and ideological considerations whereby users of the construction can identify themselves as belonging to certain subcultures or subgroups of speakers.

Keywords: language change; social identity; translation; verb complementation

1. Introduction

The present paper seeks to resolve the question of whether the Finnish language structure *rakastan puhua* ('I love to talk/speak'), which appears to be a relatively recent newcomer in the syntax of the language, is gaining ground among users of the language because of the influence of English on Finnish syntax via translation, or whether the variation between this infinitival complement structure and the more established construction *rakastan puhumista* ('I love speaking') should be accounted for by reference to some other factors. Here, then, is the syntactic essence of the two constructions in question:

(1) Rakasta-n puhua.
love-PRS.1SG talk/speak.INF
'I love to talk/speak.'

DOI 10.1075/slcs.195.12hie

(2) Rakasta-n puhumis-ta.
love-PRS.1SG talking/speaking-PART.SG
'I love talking/speaking.'

Although there clearly is considerable recent interest in various aspects of complementation in a variety of languages (see e.g. Höglund et al. 2015), it seems that the potential influence of translation on complementation via transfer has not received very much scholarly attention. And yet, it is explicitly argued by Amouzadeh and House (2010: 10) that translation is a language contact phenomenon which influences the grammatical and semantic categories of the target language as well as certain pragmatic phenomena on the target language discourse level. This then typically results in a kind of stylistic variation which did not exist in the target language previously and which is therefore triggered by some properties of the source language text.

While Amouzadeh and House explicitly deal with the relationship between English and Persian only, it appears that the phenomenon of contact-induced, translation-mediated grammatical change is observable in other languages as well, and is thus perhaps even a language universal of a sort. This, however, must remain merely a conjecture now if Amouzadeh and House (2010: 56) are correct to argue that "there are no systematic studies in contact linguistics dealing specifically with translation."

Still, it is clear that an instantiation of this phenomenon is making itself known in the Finnish language too, and that in this case it is of a relatively recent origin.

In 2011, Riitta Korhonen, one of the researchers in the Institute for the Languages of Finland (*Kotus*, < www.kotus.fi/en >) published a blog text entitled *Mehän rakastetaan puhua* ('we love to talk/speak'). The author of the blog deliberately used as the title the above syntactic structure which some readers, not very surprisingly, very soon reacted to. The reactions may be partly explainable also by reference to the fact that the structure used by Korhonen contained the so-called passive form of the verb, often used in colloquial speech:

(3) Me-hän rakaste-taan puhua.
we-surely love-PASS talk/speak
'But we love to talk/speak!'

Thus, one of the readers, Jukka Korpela, formerly associated with the Technical University (a partial predecessor of Aalto University in Helsinki, Finland) and a relatively well-known figure especially in informal and non-academic Finnish Internet discussions concerning language, commented on the blog by writing that "I assumed straight away that the title *Mehän rakastetaan puhua* was a piece of a middle-weight take-off on stupid anglicisms." In her reply, Korhonen wrote that

the formulation she used for the title was there for a reason, and that in the future such structures "will probably be neutralised and replaced by other things that people will again wonder about."

This single online discussion, despite its brevity, seems to be indicative of a change which is taking place in the Finnish complementation system. On the basis of the contribution by another participant in the above-mentioned online discussion, it turns out that a simple Google search in 2011 yielded some 3500 hits on *rakastan tehdä* ('I love to do') while in 2016 a similar search comes up with far more than 8000 hits. Thus, despite the well-known limitations and peculiarities of Google searches (cf. e.g. Brophy & Bawden 2005), it appears that Korhonen (ibid.) is correct to argue that while there are usage guides on Finnish which label the *rakastaa tehdä* ('love to do')type of structure as "somewhat alien [to present-day Finnish]" and recommend that the structure be replaced by other types of formulations at least in neutral nonfiction texts, it is also quite likely that the new structure is gradually gaining ground.

That this is indeed so is a claim which appears to be supported by some other, Google-independent data. First, by culling from the data available in the archives of the *Helsingin Sanomat*, the largest Finnish daily, one can readily ascertain that there are only a total of eight instances of the *rakastan tehdä* ('I love to do') structure in the entire archives dating from 2001 to 2013. If one replaces the verb *tehdä* ('do, make') by *puhua* ('speak, talk'), the result is even more scarce: one hit. And by using the verb *jutella* ('chat'), the result is zero. Thus, during the last few decades, the structure has entered the language with some new verbs but has not yet properly established itself as a legitimate syntactic construction with any old verb that there is in the Finnish language.

Secondly, by using collections which are far larger than the archives of a single newspaper, one can ascertain that the change is indeed a real one. For instance, in the Syntax Archives of Finnish at https://korp.csc.fi/ the collection of Finnish newspapers and magazines from the 1990s and 2000s contains almost 80 million words in more than 6.6 million sentences, and yet there are only three instances of the *rakastan puhua* ('I love to talk/speak') type of construction.

Similarly, a search on all the fourteen classical works of Finnish fiction from the 19th and 20th centuries available at the Institute for the Languages of Finland at < http://kaino.kotus.fi/korpushaku/klassikkohaku.xql > will yield a result which contains no instances of the type in question. Thus, in the 19th and 20th centuries the construction clearly had not yet established itself in the language in the genre represented by this corpus.

Further, by examining data which are culled specifically from Internet-related discussions, e.g. the Suomi24 corpus containing data from 2001 through 2016 at < https://korp.csc.fi >, one can ascertain that our variety of the construction in

Finnish is a modern invention and that one of the channels through which it is entering the Finnish language is the Internet: in this corpus there are already a total of 51 instances of the construction with the verb *puhua* ('talk, speak'), i.e. *rakastan puhua* ('I love to talk/speak'), in some 256 million sentences with approximately 2.6 billion words. Interestingly, there are only four instances of the traditional *rakastan puhumista* ('I love talking') in this corpus of less edited texts (even if there has been some moderation in place).

Why, then, is this change taking place in the language now? It is argued by Lindquist (2009: 168) that "it is one thing to show that [language] change has taken place and quite another to explain *why* change occurred," and that some of these changes in language are caused by external factors "such as fashion or language contact, when speakers of different languages mix and interact." In the case of the *rakastan tehdä* ('I love to do') structure, one such external factor seems to be the influence of English even if the impact may not be due to this foreign influence to the extent that some, especially layperson analysts, appear to think.

First, as noted by Korhonen (ibid.), there are other, distinctly domestic verbs such as *pelätä* ('fear') which already behave syntactically in the way of *rakastan puhua/tehdä/jutella* ('I love to talk/speak/do/chat') and whose syntactic behaviour in the language on the pattern in question was established a long time ago. Thus, using again a simple Google search on e.g. *(hän) pelkää puhua* ('he/she fears/is afraid to talk'), one can readily ascertain that the structure with this verb is fully productive in Finnish texts – and remarkably frequent as compared to the number of instances where the structure contains the verb *rakastaa* ('love'). Therefore, it seems reasonable to argue that the emergence of the syntactic innovation with *rakastaa* ('love') is not explainable solely and perhaps not even mainly by reference to language contact between Finnish and English.

Secondly, if the verb *rakastaa* ('love') is replaced by the Swedish-based cognate *tykätä* ('like', from Swedish *tycka*), native speakers seem to experience no oddity at all:

> Työkaveri melkein suuttuu, kun emme tykkää matkustella ulkomailla.
> '[My] colleague almost gets angry about us not liking to travel abroad.'
> (Kaleva [a Finnish daily], 13 January 2016 at < http://www.kaleva.fi/juttutupa/oma-elama/tyotoveri-melkein-suuttuu-kun-emme-tykkaa-matkustella-ulkomailla/4633111 >)

On the basis of such evidence, one can then argue that the Finnish complementation system is undergoing a change which is offering a choice to those using the language even if the choice is not necessarily a stylistic one in any simple sense (cf. Amouzadeh & House 2010: 54). Specifically, we may suggest that the choice is as follows: if a speaker chooses to use the established *rakastan puhumista* ('I love talking/speaking') structure, there may be an implicit concession on the part of

the speaker to abide by the generally accepted rules of usage and mingle with the crowd so as not to make other speakers frown upon his or her linguistic behaviour. By contrast, by using the distinctly innovative *rakastan puhua* ('I love to talk/speak') structure the speaker can indicate to his or her interlocutors that a recognisably marked variety of speech is being selected and that the speaker is therefore – more or less intentionally – making his or her interlocutors identify the speaker as a member of a specific group of language users or of a specific subculture (cf. Lindquist's (2009: 168) contention that "there are marked differences between the way[s] different age groups speak"). This situation in turn may come about at least for two different reasons.

First, it may be that it is a question of what is considered fashionable in language and that the choice is thus one which is made on the basis of certain individual sociolinguistic value judgements.[1] Secondly, it may also be a question of the somewhat wider issues of language power and language ideology (cf. e.g. Fairclough 2001, 2007). Let us consider the two cases each in turn.

2. Language and fashion

The Guardian on 6 May 2014 featured an article by David Shariatmadari where language change was assessed as follows:

> People get upset about language change. The innocent use of the word "less" where a shrinking faction insist on "fewer" is enough to earn a bad grammar award nomination and spark a national debate. If you use "enormity" to help you describe something enormous, like Barack Obama did, you'll have the likes of Simon Heffer on your back, and a few others besides.
>
> No matter how often we are told that change is perfectly natural, no matter how many examples of standard usage that our ancestors would have regarded as dangerously sloppy we are shown, the row continues.

The text of the article is yet another indication of how important it is to many if not most language users that language offer as few surprises as possible. This is a point of view which is fully understandable in light of the fact that surprises are like exceptions: they cannot be handled by general rules but rather call for some special measures, which in turn means that, with every exception, whether it comes in the form of a linguistic surprise or in some other type of disguise, an extra cognitive load is placed on the brain, and that the cognitive processes we use for

1. For a recent concise account of sociolinguistic work in Finland and other Nordic countries, see Boyd (2010). For a Finnish-language review of sociolinguistics in Finland from the 1970s to the present, see Paunonen (2009).

understanding language and thus making sense of the world consume our mental resources to an exceptionally high extent. This situation in turn can be interpreted as a type of threat to our well-being (cf. Hietaranta 2014), which is why the brain needs to resolve the situation quickly enough so as not to panic and make a bad call on what should be done next.

By contrast, expressions which are frequently used and meet our expectations about language use are processed with relative ease and do not cause any excessive load on the brain. However, a linguistic innovation is always initially an exception to some extent, but cases which are based, in one way or another, on existing linguistic practices are handled more easily than other types of innovations such as unmodified loanwords, for instance. Thus, there is no problem ascertaining that e.g. the English *college* (sweater) is still quite difficult for many native speakers of Finnish, the Finnish language containing neither the letter *c* nor the letter *g* in its domestic alphabet, which is why it is not at all uncommon to see spellings such as *gollege* or *collece* even in upmarket establishments.

Against this background, consider now the case of the *rakastan puhua* ('I love to speak') construction. Since the syntax of the language did not previously allow for structures such as *rakastan puhua* or *rakastan uida* ('I love to swim'), which the Institute for the Languages of Finland also regarded, structurally, as "borderline cases" in general language even as late as 2015 (see < http://www.kielikello.fi/index.php?mid=2&pid=11&aid=2802 >), it is not surprising to see language users react against the appearance of new verbs in the structure. Yet, since the construction is not entirely unheard of, the syntactic form in question having existed in the language with certain verbs for a long time, the reactions are, understandably, not unconditional rejections but rather expressions of surprise or as we have seen, occasionally, of even consternation.

It is not clear whether there is any semantic difference between the newer *rakastan puhua* ('I love to talk/speak') construction and the more traditional *rakastan puhumista* ('I love talking/speaking'). To resolve that question, we would need a much more extensive and a much more recent set of data than is currently available anywhere. On the other hand, the situation is in no way unique as regards studies in language change in general: in his corpus study of the variants of the sentential complementation to the verb *submit*, Rudanko (2010: 64) explicitly notes the following:

> It was also observed that the two variant types of sentential complement were found in one and the same text by one and the same author or in texts by one author that were only a few years apart. It is not easy to find a clear and consistent semantic distinction to explain the alternation.

Thus, while there may be a sociolinguistic, user identity related issue involved at least in some instances of the selection of the variant under discussion here, it

is not clear that it is the sociolinguistic aspect which is the decisive factor in the process. In fact, it seems that if we wish to look for sociolinguistic explanations for such an emergent change, we may do wisely to take a wider and more versatile look at the position of the whole of the English language as a tool for global communication in general and as its dominant position as the first foreign language in the Finnish education system. Let us therefore turn our attention to the question of the relationship between language, power, and ideology.[2]

3. Language change: An ideological note

That translation does have certain effects on how texts are interpreted and understood in their new forms in the target languages is a claim which is supported by a considerable body of empirical evidence. Thus Kuo and Nakamura (2005: 410), for example, conclude on the basis of their analysis of two translated Chinese news articles that "even though both news articles are translated from an identical English text, noticeable differences are found with respect to editorial deletions and additions, syntactic and lexical variations, as well as stylistic differences in paragraph/thematic combinations." Significantly, the authors then further explicitly argue (*ibid.*) that the "marked structural choices made by these two newspapers are not arbitrary but are well motivated by their underlying ideologies."

Similarly, when Valerio (2013: 986) analyses the Italian short novel *Pinocchio* and its three translations in English, Russian and Turkish, she concludes that the translation choices made are in fact determined to a nontrivial extent by the specific cultural backgrounds and ruling powers involved, and notes, explicitly, that "ideology cannot be avoided and translated texts are embedded with it… ."

As regards our present case of the *rakastan puhua* ('I love to talk/speak') construction, it therefore seems reasonable to conclude that another explanation for the emergence of the innovation may lie in the direction of language ideology and cultural power. Specifically, I suggest that the new usage has emerged because of the ideology and power attached to the use of the English language in Finland. This contention is based on the fact that English is the dominant language in much of current educational and social life in the country. For one thing, English is the first foreign language selected by more than 90% of pupils (or rather, by

2. I am following here van Dijk (2006: 115) and use the term *ideology* in the following sense: "As 'systems of ideas', ideologies are sociocognitively defined as shared representations of social groups, and more specifically as the 'axiomatic' princip[l]es of such representations. As the basis of a social group's self-image, ideologies organize its identity, actions, aims, norms and values, and resources as well as its relations to other social groups."

their parents) at comprehensive schools. It is then not very surprising that especially the younger generations seem to accept the *rakastan puhua* ('I love to talk/speak') structure quite readily. Secondly, English is the language from which most foreign-language books are translated into Finnish (see e.g. http://www.stat.fi/tup/suomi90/huhtikuu.html), which is why many aspects of the Anglophone cultures tend to seep into Finnish society and life over time. Thirdly, it is English-language films which are the most popular in Finnish cinemas: among the all-time top 10 films in Finnish cinemas, there are only three foreign films, all of them English-language films from the USA.[3] In consequence, it is no real surprise that traces of various linguistic practices from the English language are gradually appearing more and more frequently in Finnish, not only in the form of loanwords such as *leggin(g)sit*, *CD* and *selfie* but also perhaps structurally. That language users occasionally do face a real problem with these newcomers is indicated here by the fact that the Finnish loanword equivalent to the English *leggings* is written either with or without a *g* in front of the plural nominative ending *-it*. The two spellings are roughly of equal frequency as of this writing (see Google, s.v. *leggingsit*, *legginsit*).

It is not, however, fully clear in any great detail how and why English is having such a distinct impact on the Finnish language and Finnish culture at present. There are many instances where Finnish has chosen to ignore such foreign influence as time has gone by (cf. e.g. *puhelin* 'telephone', from *puhua* 'speak, talk'), which may of course be the fate of some of the present-day loanwords as well; they may be replaced by more domestic equivalents in time.

One explanation which one might suggest for the upper hand the English language is having these days within Finnish culture, including language use, is the fact that the influence English is having on other languages and cultures through translation is of a very subtle kind. That is, since culture is typically encoded in language in less than explicit ways, speakers of a language who borrow items from another language – especially if they are linguistically naïve, lay users of the language – are seldom capable of detecting what kinds of values and worldviews are attached to the borrowed items. Therefore, it is even more likely that, in those cases where there are no obvious concomitant values in the donating language items involved, the borrowers will be even more willing to accept the innovation when the origin of the innovation is not clearly visible. This, it seems to me, may be one reason behind the willingness of some Finnish speakers to accept the new *rakastan puhua* ('I love to talk/speak') variant into the language without further ado.

Another aspect of the issue is the use of English as a lingua franca in much of present-day international communication. Because of this status, English is

3. <https://fi.wikipedia.org/wiki/Luettelo_Suomen_katsotuimmista_elokuvista >

understandably gaining more and more weight as a means of communication, which in turn may well add to the relative weight and acceptability of whatever is transferred from English into other languages and cultures. Such a situation may then lead some language users to resist language change – on the syntactic or some other level – because of the assumption that such a mixing of languages and cultures may lead to misunderstandings and other communication problems when the norms of two language systems are being used more or less simultaneously. This assumption, however, seems to be a problematic one.

Kaur (2011) reports on a study where the effects of English as a lingua franca on understandability in communication were investigated, and concludes that "none of the misunderstandings in the data can be traced to differences in the participants' cultural background." Given that Kaur is correct to draw this conclusion, which has also been drawn independently on different sets of data by House (1999) and by Mauranen (2006), one may argue that the status of English as a lingua franca probably does very little to scare language users of other languages as a potential source of misunderstanding in particular and of communication problems in general. Given that this is so, it becomes even more understandable that users of other languages should assume favourable attitudes towards English-based innovations in the form of language change: if the influence of English is not distinctly harmful, why not use yet another linguistic practice which clearly works for English and which seems to fit our language system too?

Speakers of other languages are of course hardly ever aware of the fact that English is not demonstrably harmful to or even affecting their languages and cultures. However, what matters here is another issue, viz. the fact that as long as a borrowed item or linguistic practice works for a language, its users see no need to reject the loan. If the new or, as in our case, extended practice enables language users to achieve their communication goals, many of them will see no reason to resist the change. Only those who wish to rely on proven methods are likely to regard innovations as potential threats, and will start expressing their doubts about the need for the change and its viability. Here, issues of cultural and other types of prestige, ideology and language power may start surfacing especially in cases where speakers find it important to align their communication habits with the expectations of their interlocutors as argued by Weatherholtz et al. (2014).

4. Speaker-alignment as recipient design: Domestication and foreignisation generalised

Weatherholtz et al. (2014) is a study of how language users align their linguistic behaviours when they interact with their interlocutors. The study is of special

significance to our present concerns in that the authors explicitly note that while there was "an overall alignment effect across social conditions and independent of social perceptions," it was also the case that "the degree of alignment depends on participants' perceptions of others, participants' individual tendencies, and the linguistic structures in the ambient environment that are available to align (or antialign) with" (Weatherholtz et al. 2014: 387, 412). It is possible to use these observations to account for the limited occurrence of the *rakastan puhua* ('I love to talk/speak') structure in Finnish by viewing the authors' treatment of the subject as yet another form of what Mustajoki (2012: 227–231) calls recipient design in his analysis of risks and causes of miscommunication, i.e. "the speaker's strategy of adapting her/his communicative behaviour to the recipient." That language users generally seem to make use of this possibility of adapting their linguistic and communicative behaviour towards other language users is explainable by reference to a number of factors. Here, I will consider only three of them.

Consider first the observation that there were so few instances of the structure in the data examined, and that search engines such as Google nevertheless now find such a multitude of cases online. Why should this be so? I suggest that the explanation for this fact is that newspaper texts are designed to be palatable by any reader and for that reason do not need to be aligned in any particular way at all since there is no particular segment of the readership which the texts should or even could be aligned with. That is, newspaper texts should be approachable by virtually anyone, which means that they should be, in terms of their textual design, including their syntax, as neutral and in that way as general as possible. For this reason, a syntactic form which is distinctly marked rather than unmarked is unlikely to appear in such general-purpose texts. Thus, there is no need for any particular type of recipient design in terms of specific forms of alignment as far as newspaper texts are concerned.

By contrast, online discussions typically consist of texts written by language users who belong to groups where the members share some common interests and therefore not only identify themselves with specific language behaviours as such but expect certain types of linguistic interaction on the part of their fellow members. To meet such expectations, group members need to indicate to their peers by textual means, including the syntactic structures used in text production, that the expectations are being duly noted and are being fulfilled in that certain types of structures are being resorted to. It is then for this reason that language users in online discussion forums and other similar settings frequently make use of syntactic structures and other linguistic devices whose frequency of occurrence in general-language texts is notably lower. In short, in online discussions there are well-founded reasons for aligning the texts produced so that they will meet

specific recipient design requirements in terms of identifiability, subculture, and ideology or worldview.

Secondly, consider the influence of English on Finnish and other languages through translation. Again, it seems reasonable to assume that general-language and especially general-purpose texts, a major part of which is constituted by translated literary works, are texts where the authors need to reach audiences which cannot be specified by any small number of criteria. For instance, fiction texts may be and typically are read by people who do not constitute any kind of specific readership or other community or group at all. For this reason, such texts, also in translation, should not contain any passages where the text is insufficiently general in terms of its design – again including the syntactic structures used. Otherwise the text may prove insurmountable or uninteresting to some readers because of its unintentional markedness for one or another linguistic property.

Yet, as regards translation, we should also note that translations are often either domesticated or foreignised, that is, a text can be translated either "in a transparent, fluent, 'invisible' style in order to minimize the foreignness of the TT [target text]" (Munday 2012: 218), or, alternatively, the translator may opt for "a non-fluent, estranging or heterogeneous translation style designed to make visible the presence of the translator and to highlight the foreign identity of the ST [source text]" (Munday 2012: 219). In line with this distinction, I suggest that just as a translator may choose to translate in either a domesticating or a foreignising manner, language users in general may choose either to align or antialign their spoken and written texts with the expectations of their known or anticipated recipients. Language users do this either to identify themselves with particular subgroups of their language communities (with a particular age group or with a group propagating a particular ideology or worldview, for instance), which is what happens when language users do decide to align their communicative products in terms of specific recipient designs, or else they wish to indicate to their interlocutors that they prefer to dissociate themselves from particular subgroups or subcultures and be rather part of the great majority. In the latter case they will behave, linguistically and probably also in other respects, in inconspicuous or at least less conspicuous ways and thus in general act in ways which will enable them not to attract too much attention to themselves.

Consider now how the notion of recipient design could be used to account for the emergence of the *rakastan puhua* ('I love to talk/speak') construction in Finnish. If, as e.g. the Archives of Finnish Syntax data seem to indicate (only three instances of the construction in the corpora consisting of Finnish newspaper and magazine texts from the 1990s and 2000s, close to a total of 80 million words), the construction is a relatively late addition to the grammar of Finnish, it is unlikely that those speakers of the language who have used the structure with other,

established verbs or verb groups for decades could be the origin of the innovation; rather, it seems more likely that that the change was initiated by users who belong to a younger age group and whose linguistic habits were therefore not yet as fixed and rigid as those of the older generations, and who were, accordingly, willing to experiment with and ultimately accept a new syntactic norm for the language more readily than members of the older generations. Furthermore, speakers with such a more flexible attitude towards linguistic norms could also benefit from yet another property of recipient design, viz. the intentional decision to ignore monitoring one's own language use and the associated recipient design.

Specifically, as noted by Mustajoki (2012: 228), there are certain cases or situations which offer little if any justification for monitoring one's linguistic or communicative behaviour. Typically, if we are interacting with people we know intimately (family members, close colleagues) or if even we are socialising in a pub with people we do not know intimately but the atmosphere is sufficiently relaxed, there seems to be little reason to monitor one's own behaviour explicitly or very closely. Instead, we tend to assume that the requirements with regard to recipient design may be temporarily relaxed at least somewhat, and that there will still be no misunderstanding as to what we are saying or which group(s) of language users, if any, we belong to. Consider now how such an apparently justified attitude of relaxation towards recipient design may account for the recent emergence of *rakastan puhua* ('I love to talk/speak') in Finnish.

People participating in online discussions are usually aware of what kind of linguistic activity they are involved in and what is expected of them in such settings. If someone is not familiar with the rules and regularities of the forum he or she is interested in, the person often starts the first post by writing "I am new here, so…" or by providing some other similar indication of his or her background and relation to the group the forum is intended for. In such a context, users then very soon develop a feel for what kinds of issues must be specified explicitly whenever they publish a new post and what can be assumed to be self-evident or at least what can be legitimately implied. People typically behave in this way because, as noted by Mustajoki (2012: 228), people tend to avoid extra cognitive and/or physical effort. If run-of-the-mill behaviour, including linguistic behaviour, is enough, that is what most people will settle for, relying on common ground and implying rather than explicitly indicating some of the relevant aspects of their messages. Now, while it is probably true that the common ground fallacy, that is, an unjustifiably large portion of trust in mutual understanding on the basis of shared cultural background(s), does exist (cf. e.g. Clark 1996: 111), it is also the case that members of special interest groups sharing interests and values can often rely on each other's ability to interpret messages correctly, that is, in the intended manner. This, I suggest, is to a notable extent so because the members in such groups operate in

their respective contexts with limited sets of linguistic data and communicative practices, and can therefore safely assume that the community will react to any communicative deviation soon enough to prevent or at least correct misunderstandings. But this is not all there is to being an insider: once you have become accustomed to what it means to be a member of a particular group or subculture, you can use many of the group's or culture's existing practices as shortcuts to send messages to other members of your community, and because you are using shortcuts, you are saving cognitive effort: since you can assume, on the basis of the nature of such a group, that group pressure will keep any excessive communicative behaviour at bay, it follows that you do not need to spell out on each individual occasion of language use how you wish other members to perceive you in relation to them or to your shared values; rather, you can achieve your goal(s) by using suitable structures and other fragments of language when communicating with your fellow members. Thus, it seems to me, one explanation for the emergence of the *rakastan puhua* ('I love to talk/speak') construction is not some speakers' desire to set themselves apart from the main body of Finnish speakers as such but rather the desire to indicate to other members of their group, which is probably constituted by a cohort of relatively young speakers of Finnish, that "we are also members of the group you belong to, we share some common goals and values."

Assuming that the members of the group(s) who are more or less regularly using the construction belong to the younger generation(s), it is also plausible to think that the members of these age groups show the least resistance to such innovations because they have had English as their first or main foreign language at schools and have been very much exposed to Anglophone cultures in their daily lives (British and American television programmes and films, popular music, the computer world, the Internet, magazines and books translated from English).[4] That is, if such young(ish) language users encounter in their mother tongue a new construction or expression which has a seemingly similar counterpart in English, it is all the more likely that they will not frown upon the newcomer but will instead note that the stranger is somewhat different from what they are used to seeing but at the same time also similar to something they have already witnessed elsewhere so that the construction is indeed, in consequence, in time "neutralised and replaced by other things that people will again wonder about" (Korhonen 2011).

There are, of course, also other ways in which translations may affect target language systems. For example, as argued by Taylor (2008), Old English was influenced by translations from Latin both directly and indirectly when different

4. For an account of how important age is as an explanatory factor for the Internet usage of different types of people in general, see Brandtzæg et al. (2011).

translators applied different translation strategies to the texts they were working on. In particular, biblical and nonbiblical texts seem to have been treated somewhat differently in translation, which is an interesting observation in view of the fact that the structure under discussion here, *rakastan puhua* ('I love to talk/speak'), appears to be limited in distribution to less formal contexts (cf. Korhonen 2011). Given that this is so, we may consider the current distribution of the structure as yet another indication of how linguistic changes progress over time in ways which are conditioned by both linguistic and sociological factors.

5. Final remarks: Language variation and language change diffusion

A final issue that I wish to consider here in relation to the emergence of the *rakastan puhua* ('I love to talk/speak') construction in Finnish syntax is the question regarding the extent to which speakers may and do participate in and thereby expedite language change over time. The issue is addressed e.g. by Nevalainen et al. (2011), who specifically examine the phenomenon in light of the hypothesis that in any particular period the majority of speakers are neither conservative nor progressive with regard to their attitudes towards language change.

That this question is worth examining in detail is so because of the fact that it has not always been very clear what kinds of factors contribute to the emergence of change in language. Thus it is argued by Kristiansen et al. (2005: 9) that

> if we want to answer the *why* question of variation and change - why does it happen? - we do think that there is a fundamental distinction to be drawn between socio-psychological or subjective factors and processes, on the one hand, and all the others - objective factors and processes - on the other hand. In other words, we believe in and wish to stress the relative independence and importance of subjective factors in relation to objective factors and their interaction in the processes of LVC [language variation and change].

As I have argued above, the fact that the *rakastan puhua* ('I love to talk/speak') construction has emerged in Finnish should be attributed to some socio-psychological factors rather than to some system-wide pressures or processes – unlike e.g. the desire of many usage guides or authorities on language use to change some practices so that the practices in question will become more compatible with more or less general rules of language instead of standing out as exceptions, one illustrative case in point being the Finnish noun *sovellus* ('application', also in the sense of 'software application'). Previously, the form of the noun used was *sovellutus* (which one may still come across occasionally in texts representing more than one text type, s.v. *sovellutus* in *Syntax Archives of Finnish*). That the recommended

form is now the shorter variant is due to the fact that this form is a regular derivative of the stem verb *soveltaa* ('to apply') and thus exhibits behaviour which is fully regular and for that reason predictable on the basis of a general derivational rule: in addition to *sovellus*, Finnish also has *sävellys* ('composition', from *säveltää* 'to compose music'), *kavallus* ('embezzlement', from *kavaltaa* 'to embezzle'), and *vihellys* ('whistle' (noun), from *viheltää* 'to whistle'), among others. For these morphological changes, it is thus the pressure of the system which makes us prefer the shorter *sovellus* to the older and longer *sovellutus* whereas there is no such pressure from the language system itself and accordingly no objective justification for changing the Finnish complementation system to accept *rakastan puhua* ('I love to talk/speak') as a fully-fledged member of the structural paradigm of the language.

The above view of the status of the construction with regard to the change noted in the Finnish complementation system is of distinct theoretical interest also in that it is argued by Nevalainen et al. (2011: 5), in line with Labov (2001: 28–29), that "people vary more in their use of changing abstract structural patterns because these processes are … less likely to be associated with social evaluation than lexical and phonological changes…" In consonance with this abstract structural pattern hypothesis, it is a structure of the Finnish grammar which is undergoing a change here.

Consequently, I suggest that the change we are discussing here is taking place also because language users are only less likely to initiate and approve language change on the lexical and phonological levels than on the structural levels rather than being completely opposed to such structural change. That is, while it may be the case that language change is encountered more frequently on the lexical and phonological levels than in other parts of the system, there may also arise circumstances under which users may opt for a structural change – for reasons such as those discussed above.

Of further interest is the observation by Nevalainen et al. (2011: 26) that "it is possible to identify leaders, laggards, and in-betweens in the diffusion of six changes in Early Modern English." Assuming that it is possible to extrapolate from this conclusion to language change more generally, it becomes possible to argue that the recent *rakastan puhua* ('I love to talk/speak') variant is now there in the Finnish complementation system because it enables its users to identify themselves (perhaps unconsciously) as leaders of some sort: those opting for this variant may not be overly progressive in the sense of Nevalainen et al. (2011), but what is more important here is the fact that such usage may signal to other speakers of the language a certain willingness to admit that English is a reputable source of influence not only in terms of the various Anglophone cultures transmitted through language but also as regards the possibility of finding new structural means of expression under certain sociologically motivated communicative circumstances.

Finally, I wish to argue that the diffusion of the change we have been examining here is tied to a nontrivial extent to the fact that while linguistic leadership in change may well be connected with social networks (Nevalainen et al. 2011: 26), it is also the case that "in the incipient phase, leaders may be found among geographically mobile people, who may be expected to have a great many weak links, a characteristic that … promotes the diffusion of linguistic changes" (ibid.). Specifically, I suggest that in modern times many of the weak links referred to in the above quotation are constituted by online contacts over the Internet, and that because the Internet is such an overwhelmingly powerful channel of influence, a remarkable portion of this influence is dribbling to a sufficiently significant number of language users, sufficiently many of whom are ready to accept the new usage as a fully respectable means of expression. That is, in our times it is no longer people or language users themselves which are mobile but rather the channel through which communication, with all its potential or actual novelties, is being transmitted. For this reason, it is now possible to have considerable diffusion for virtually any kind of linguistic innovation even at the initial stage of the change, a further contributing factor being the fact that a sizeable portion of the data moving on the Internet is never censored or even monitored on linguistic grounds: while there are moderators on many forums, in the capacity of naïve, layperson language users, they are often neither capable of nor interested in assessing the posts submitted for publication on the basis of the posts' linguistic or communicative properties. Accordingly, even innovations which are more or less clearly of foreign origin may start spreading in a language system so that there will soon be no way of stopping them.

Acknowledgements

I wish to thank the two anonymous reviewers of this paper whose comments on the original version of the text made me revise and expand considerably on especially the part of the argumentation which concerns the data utilised and who thus enabled me to improve on the original version of the present contribution to a notable extent. Any remaining inadequacies are of course my sole responsibility.

References

Amouzadeh, Mohammad & House, Juliane. 2010. Translation as a language contact phenomenon: The case of English and Persian passives. *Languages in Contrast* 10(1): 54–75. doi: 10.1075/lic.10.1.03amo

Boyd, Sally. 2010. Sociolinguistics in the Nordic region. In *The Routledge Handbook of Sociolinguistics around the World*, Martin J. Ball (ed.), 281–296. London: Routledge.

Brandtzæg, Petter Bae, Heim, Jan & Karahasanović, Amela. 2011. Understanding the new digital divide – A typology of Internet users in Europe. *International Journal of Human-Computer Studies* 69: 123–138. doi:10.1016/j.ijhcs.2010.11.004

Brophy, Jan & Bawden, David. 2005. Is Google enough? Comparison of an internet search engine with academic library resources. *Aslib Proceedings: New Information Perspectives* 57(6): 498–512. doi:10.1108/00012530510634235

Clark, Herbert H. 1996. *Using Language*. Cambridge: CUP. doi:10.1017/CBO9780511620539

Fairclough, Norman. 2001. *Language and Power*, 2nd edn. Harlow: Longman.

Fairclough, Norman. 2007. *Language and Globalization*. London: Routledge.

Hietaranta, Pertti. 2014. Viewing translation problems as threats: a cognitive approach to misunderstandings in translation. Paper presented at the international *Translation in Transition: Between Cognition, Computing and Technology Conference*, Copenhagen, 30–31 January.

House, Juliane. 1999. Misunderstanding in intercultural communication: Interactions in English as a lingua franca and the myth of mutual intelligibility. In *Teaching and Learning English as a Global Language*, Claus Gnutzmann (ed.), 73–89. Tübingen: Stauffenburg.

House, Juliane. 2003. English as lingua franca: A threat to multilingualism? *Journal of Sociolinguistics* 7(4): 556–578. doi:10.1111/j.1467-9841.2003.00242.x

Höglund, Mikko, Rickman, Paul, Rudanko, Juhani & Havu, Jukka. 2015. *Perspectives on Complementation: Structure, Variation and Boundaries*. Houndmills: Palgrave Macmillan.

Institute for the Languages of Finland: corpora and other material. <https://www.kotus.fi/en/corpora_and_other_material> (2 May 2017).

Kaur, Jagdish. 2011. Intercultural communication in English a lingua franca: Some sources of misunderstanding. *Intercultural Pragmatics* 8(1): 93–116. doi:10.1515/IPRG.2011.004

Korhonen, Riitta. 2011. Mehän rakastetaan puhua! ('But we love to speak!'). Blog. <http://www.kotus.fi/nyt/kotus-blogi/blogiarkisto/riitta_korhonen/mehan_rakastetaan_puhua!.8394.blog> (30 June 2016).

Korhonen, Riitta & Maamies, Sari. 2015. Kieliopillisten vaihtoehtojen välissä – tiukkoja ja väljiä normeja. ('Between grammatical choices – strict and loose norms'). *Kielikello* 2(2015). <http://www.kielikello.fi/index.php?mid=2&pid=11&aid=2802> (30 April 2017).

Kristiansen, Tore, Garrett, Peter & Coupland, Nikolas. 2005. Introducing subjectivities in language variation and change. *Acta Linguistica Hafniensia* 37(1): 9–35. doi:10.1080/03740463.2005.10416081

Kuo, Sai-Hua & Nakamura, Mari. 2005. Translation or transformation? A case study of language and ideology in the Taiwanese press. *Discourse & Society* 16(3): 393–417. doi:10.1177/0957926505051172

Labov, William. 2001. *Principles of Linguistic Change*, Vol. 2: Social factors. Oxford: Blackwell.

Lindquist, Hans. 2009. *Corpus Linguistics and the Description of English*. Edinburgh: EUP.

Mauranen, Anna. 2006. Signaling and preventing misunderstanding in English as lingua franca communication. *International Journal of the Sociology of Language* 177: 123–150.

Munday, Jeremy. 2012. *Introducing Translation Studies: Theories and Applications*. London: Routledge.

Mustajoki, Arto. 2012. A speaker-oriented multidimensional approach to risks and causes of miscommunication. *Language and Dialogue* 2(2): 216–243. doi:10.1075/ld.2.2.03mus

Nevalainen, Terttu, Raumolin-Brunberg, Helena & Mannila, Heikki. 2011. The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change* 23: 1–43. doi:10.1017/S0954394510000207

Paunonen, Heikki. 2009. Suomalaisen sosiolingvistiikan ja kielisosiologian näkymiä. *Virittäjä* 113: 557–570.

Rudanko, Juhani. 2010. Change and variation in complement selection: a case study from recent English, with evidence from large corpora. In *Corpus-linguistic Applications: Current Studies, New Directions*, Stefan T. Gries, Stefanie Wulff & Mark Davies (eds), 47–66. Amsterdam: Rodopi.

Syntax Archives of Finnish. <https://korp.csc.fi/#/lang=fi> (2 May 2017).

Taylor, Ann. 2008. Contact effects of translation: Distinguishing two kinds of influence in Old English. *Language Variation and Change* 20: 341–365. doi:10.1017/S0954394508000100

Valerio, Anna. 2013. Translation and ideology: A critical reading. *Social and Behavioral Sciences* 70: 986–996.

van Dijk, Teun A. 2006. Ideology and discourse analysis. *Journal of Political Ideologies* 11(2): 115–140. doi:10.1080/13569310600687908

Weatherholtz, Kodi, Campbell-Kibler, Kathryn & Jaeger, T. Florian. 2014. Socially-mediated syntactic alignment. *Language Variation and Change* 26: 387–420. doi:10.1017/S0954394514000155

Author index

Subject index

This book is a collection of eleven research articles which altogether serve as a contribution to the study of verb complementation and other constructions, an area of investigation which bridges observations on the spectrum of lexico-grammar, syntax, and semantics. In terms of methodological approaches and the types of linguistic patterns examined, the chapters cast light on the subject from a variety of perspectives, and the volume is structured in a way that groups the various perspectives under three main themes according to their main focus and/or methodological approaches, namely: the semantic and functional descriptions of constructions; the investigation into the distribution of complementation patterns; and the study of innovative patterns in ESL contexts and languages other than English. All chapters in this volume employ data from large electronic corpora where possible – the BNC, COCA, COHA, GloWbE, NOW, and newly compiled corpora representing regional varieties of English.

ISBN 978 90 272 0054 9

JOHN BENJAMINS PUBLISHING COMPANY