Constructional Approaches
to Language 21

Grammaticalization meets Construction Grammar

edited by
Evie Coussé
Peter Andersson
Joel Olofsson

John Benjamins Publishing Company

Grammaticalization meets Construction Grammar

Constructional Approaches to Language

ISSN 1573-594X

The series brings together research conducted within different constructional models and makes them available to scholars and students working in this and other related fields.

The topics range from descriptions of grammatical phenomena in different languages to theoretical issues concerning language acquisition, language change, and language use. The foundation of constructional research is provided by the model known as Construction Grammar (including Frame Semantics). The book series publishes studies in which this model is developed in new directions and extended through alternative approaches. Such approaches include cognitive linguistics, conceptual semantics, interaction and discourse, as well as typologically motivated alternatives, with implications both for constructional theories and for their applications in related fields such as communication studies, computational linguistics, AI, neurology, psychology, sociology, and anthropology.

This peer reviewed series is committed to innovative research and will include monographs, thematic collections of articles, and introductory textbooks.

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

Volume 21

Grammaticalization meets Construction Grammar
Edited by Evie Coussé, Peter Andersson and Joel Olofsson

Grammaticalization meets Construction Grammar

Edited by

Evie Coussé
Peter Andersson
Joel Olofsson
University of Gothenburg

John Benjamins Publishing Company
Amsterdam / Philadelphia

DOI 10.1075/cal.21

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2018002594 (PRINT) / 2018005983 (E-BOOK)

ISBN 978 90 272 0064 8 (HB)
ISBN 978 90 272 6416 9 (E-BOOK)

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

Table of contents

Theoretical issues

Grammaticalization meets Construction Grammar: Opportunities, challenges and potential incompatibilities **3**
Evie Coussé, Peter Anderson and Joel Olofsson

Three open questions in Diachronic Construction Grammar **21**
Martin Hilpert

Verb constructions

Grammaticalisation cut short: A diachronic constructional view on English posture verbs **43**
Mégane Lesuisse and Maarten Lemmens

Pseudocoordination in Norwegian: Degrees of grammaticalization and constructional variants **75**
Torodd Kinn

Evaluating grammaticalization and constructional accounts: The development of the inchoative construction with *put* verbs in Spanish **107**
Renata Enghels and Marie Comer

Noun constructions

Reduction or expansion? *A bit of* both: A case study on the development of German degree modifiers **137**
Jakob Neels and Stefan Hartmann

Type frequency, productivity and schematicity in the evolution of the Latin *secundum* NP construction **169**
Caterina Guardamagna

Constructions at sentence level

The development of the conditional *caso* construction in Spanish **205**
Anton Granvik

Constructionalization areas: The case of negation in Manchu 241
Andreas Hölzl

A Radical Construction Grammar approach to construction split in the diachrony of the spatial particles of Ancient Greek: Some theoretical preliminaries 277
Emanuel Karlsson

Construction index 312

Index 313

Theoretical issues

Grammaticalization meets Construction Grammar

Opportunities, challenges and potential incompatibilities

Evie Coussé, Peter Andersson & Joel Olofsson
University of Gothenburg

1. Introduction

Grammaticalization research has in the last decade highlighted the notion of constructions. Hopper & Traugott (2003, p. 1) in their classical definition of grammaticalization point out that that not only words but also constructions, i.e. sequences of words, can undergo grammaticalization.

> As a term referring to a research framework, "grammaticalization" refers to that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain contexts to serve grammatical functions or how grammatical items develop new grammatical functions.

Himmelmann (2004, p. 31) in turn argues that a grammaticalizing element (he uses the synonym term a 'grammaticizing' element) should not be considered in isolation but rather in its syntagmatic context, i.e. the construction it occurs in.

> Strictly speaking, it is never just the grammaticizing element that undergoes grammaticization. Instead, it is the grammaticizing element in its *syntagmatic context* which is grammaticized. That is, the unit to which grammaticization properly applies are *constructions*, not isolated lexical items.

In the wake of this work, efforts have been made in grammaticalization research to more precisely articulate the largely pretheoretical notion of construction in the theoretical framework of construction grammar. Confronting grammaticalization research with the framework of construction grammar also brought up the question how grammaticalization relates to constructional change in general (Noël, 2007; Gisborne & Patten, 2011; Heine et al., 2016). As such, grammaticalization research increasingly interacts and converges with the emerging field of

https://doi.org/10.1075/cal.21.c1

diachronic construction grammar (Israel, 1996; Bergs & Diewald, 2008; Barðdal et al., 2015).

This volume brings together articles that are situated at the intersection of grammaticalization research and diachronic construction grammar. All articles share an interest in integrating insights from grammaticalization research and construction grammar in order to advance our understanding of empirical cases of grammaticalization. This introduction lays the ground for the central recurring theoretical issues in the articles, addressing opportunities, challenges and potential incompatibilities in the growing interaction between grammaticalization and construction grammar. First out is the question what makes construction grammar so attractive for the study of grammaticalization phenomena (Section 2). A related question is how concepts from grammaticalization theory can be related to and integrated into construction grammar (Section 3 and 4). Some thought will also be given to the question what kind of construction grammar typically is used for the study of grammaticalization (Section 5). Finally, some critical voices in the field will be examined, addressing areas where grammaticalization and construction grammar (might) prove to be incompatible (Section 6).

2. Construction grammar comes with a focus on form and meaning

One of the attractions of construction grammar for the study of grammaticalization is its fundamental focus on both form and meaning. Constructions, the basic building blocks of construction grammar, are defined as pairings of form and meaning (Langacker, 1987; Fillmore et al., 1988; Croft, 2001; Goldberg, 1995, 2006). These pairings are not restricted to the level of words (as is the case in the traditional Saussurean sign) but extend to all levels of grammar. Grammar in this view is "constructions all the way down" (Goldberg, 2006, p. 18) ranging from specific words such as *avocado*, fully fixed idioms such as *going great guns* and partially filled idioms such as *jog* <someone's> *memory*, to more general syntactic structures such as the subject-predicate construction. Grammaticalization is known to be accompanied by both formal and semantic changes, as is witnessed by Lehmann's (1982) well-known criteria for grammaticalization, including both the phonetic erosion of a grammaticalizing item and its semantic bleaching. Construction grammar, with its systematic linking of form and meaning in constructions at all levels of grammar, urges us to consider both form and meaning, in every step of the grammaticalization process, and at all levels of grammar involved in the process.

3. Construction grammar and syntagmatic contexts

Constructions, as defined in construction grammar, have also a great potential for the syntagmatic dimension of grammaticalization. Grammaticalization theorists have since the beginning of the new millennium emphasized the importance of the specific syntagmatic context in which grammaticalization takes place (Heine, 2002; Diewald, 2002; Traugott, 2003; Bybee, 2003; Himmelmann, 2004). A classic example is the grammaticalization of the future marker *going to.* Bybee (2003, p. 146) argues that "*going to* does not grammaticalize in the construction exemplified by *I'm going to the store* but only in the construction in which a verb follows *to*, as in *I'm going to help you*". Construction grammar offers a framework to model these grammaticalization contexts in terms of form-meaning pairings.

The type of construction that is particularly relevant for the study of grammaticalization is what we could call the 'host-class construction'. It is a construction consisting of the grammaticalizing element (or string of elements as in *going to*) and "the class of elements the gram is in construction with, i.e. the host class" as Himmelmann (2004, p. 32) puts it. In more technical construction grammar terms, we are dealing with a semi-schematic construction, where the grammaticalizing element(s) constitute(s) the substantive element(s) of the construction and the emerging host-class forms an open slot or schematic position for lexical elements collocated with the grammaticalizing element(s). As such, *going to* is the substantive part of the semi-schematic construction [*going to* V] expressing future tense.

Note that there is no consensus on the fact that grammaticalization results in a host-class construction (see also Noël, 2007). Grammaticalization has been argued to include types of grammatical change such as word order changes and the development of discourse markers where no host-class can be identified. At any rate, it should not come as a surprise that the constructions studied in this volumes are of the semi-schematic kind.

Construction grammar does not only provide a model for describing the syntagmatic contexts in which grammaticalization takes place. It also helps us come into grips with the contextual changes that are known to accompany grammaticalization. Lehmann (1982) has identified three types of changes in the syntagmatic dimension that come with grammaticalization: decrease of structural scope (condensation), increase of bondeness (coalescence), and decrease in syntagmatic variability (fixation). The first two of these changes can directly be related to the association process between the grammaticalizing element and its host-class which Himmelmann (2004) hinted at. As the grammaticalizing element gets associated with an expanding host-class, its structural scope becomes exclusively restricted to this construction-internal class of elements (condensation), and the

collocational association between the substantive and schematic position gets strengthened (coalescence).

Traugott & Trousdale (2013, p. 100) point out that the grammaticalization model of Lehmann (1982) focuses on reduction and increased dependency. Two articles in this volume exploit these aspects of grammaticalization on the syntagmatic axis in order to uncover ongoing grammaticalization. Kinn (this volume) explores constructional variants of the pseudocoordinated construction [VP_1 *og* VP_2] in Norwegian, such as *sitte og lese* 'be reading' (literary: 'sit and read'), *gå og se* 'go and see', and *løpe rundt og leke* 'run around and play'. He argues that a high degree of grammaticalization of the first verb in the asymmetric coordination construction correlates with (a) scarcity or absence of constituents in between VP_1 and the coordinator *og*, (b) a blurring of the facilitation relation between VP_1 and VP_2, and, (c) the meaning of the coordination approaching the meaning of VP_2, thus strengthening the single-event reading of the construction. Enghels & Comer (this volume) investigate the grammaticalization of the Spanish lexical verbs *poner* and *meter* 'put' into inchoative auxiliaries. They explore similarly to Kinn (this volume) how the grammaticalization of *poner* and *meter* correlates with the presence of intervening material in the inchoative construction [NP_1 V_{refl} PREP INF]. They observe a decrease of intervening adverbial complements in inchoative constructions with *poner* and relate it to an increased syntactic incorporation of *poner* and its host-class.

Although the above criteria are related to the syntagmatic parameters of Lehmann, it is not possible to identify a one-on-one relation between criteria and parameters (see also Traugott & Trousdale, 2013, p. 123). The criteria rather exploit various reduction tendencies in host-class constructions that have been described in more general terms in diachronic construction grammar. Traugott & Trousdale (2013, p. 120–122) for instance argue that the formation of a new construction (a process they call "constructionalization") comes with a decrease of compositionality, as the result of a mismatch between the overall meaning of a construction and the meaning of its constituting parts. Traugott & Trousdale (2013, p. 122) also mention that "cumulative effects of entrenchment help explain why constructions generally (not only grammatical ones) show high degrees of fixation".

Next to grammaticalization as reduction and increased dependency, Traugott & Trousdale (2013, p. 100) highlight that grammaticalization also comes with expansion. Himmelmann (2004) has articulated two types of expansions that are directly relevant to the syntagmatic axis of grammaticalization: host-class expansion and expansion of the larger syntactic context in which the construction is used. Both characteristics have been exploited in this volume in order to uncover

ongoing grammaticalization. Enghels & Comer (this volume) investigate the relative frequency of constructional variants with a varying degree of animacy of the subjects NP_1 and a varying dynamicity of the infinitive INF. They find that the inchoative construction with *poner* expands to semantic types other than the original prototypical use with an animate NP1 and a dynamic INF. The observation confirms the claim of Coussé (2018) that context expansion in grammaticalization proceeds away from a prototype.

Context expansion in grammaticalization affects the dimension of schematicity and productivity in constructions. The schematicity of a construction pertains to the extent in which it can abstract or generalize over more specific items (Tuggy, 2007). In the process of host-class expansion, a semi-schematic construction incorporates a wider range of items in its schematic position, which leads to an overall higher degree of schematicity in the construction. Productivity is a related dimension in constructions in that it pertains to the extensibility of schematic constructions (Barðdal, 2008; Barðdal et al., 2015). Productivity can be assessed by means of frequency measures, which makes it a valuable tool in quantitative construction grammar. Two articles have exploited productivity measures in order to assess the degree of grammaticalization in their corpus material.

Lesuisse & Lemmens (this volume) investigate the degree of grammaticalization of the English posture verbs *sit*, *stand* and *lie* in adjectival copula constructions of the type [SubjectNP V SubjCompAdjP]. The productivity of these constructions is assessed by looking into specific frequency patterns of the adjectival subject complement, the expanding host-class of this construction. Lesuisse & Lemmens more specifically measure the type frequency of this open slot and its so-called productivity rate, i.e. the number of hapax legomena (types that occur only once in a corpus) divided by the total number of tokens (Baayen & Lieber, 1991). Guardamagna (this volume) similarly investigates the grammaticalization of the Latin preposition *secundum* 'according to' in the construction [*secundum* NP] measuring type frequency, token frequency, number of hapax legomena and finally Baayen's (1993) productivity index of the NP slot.

4. Construction grammar and paradigmatic relations

Construction grammar also helps us come into grips with the paradigmatic side of grammaticalization. One well-known feature of grammaticalization is that grammaticalizing elements become part of a paradigm, something that Lehmann (1982) calls 'paradigmaticization'. Take the grammaticalization of *going to*. As a lexical expression, it is part of a wide and loosely structured semantic field of motion. As it grammaticalizes and starts serving in a future tense construction,

it gets incorporated into a much smaller closed paradigm of auxiliaries expressing tense. Construction grammar approaches paradigmatic issues no differently from syntagmatic matters, i.e. by making use of its basic building blocks, constructions. However, whereas syntagmatic parameters of grammaticalization are confined to a single construction, paradigmatic issues imply a relation between two or more constructions.

Constructions are related to each other in construction grammar by means of taxonomic networks. The network, and its different levels of abstraction or schematicity, can be illustrated by means of a classic example, i.e. that of binominal degree modifiers (Figure 1). Neels & Hartmann (this volume) document the grammaticalization of German degree modifiers such as *ein bisschen* 'a bit' and *ein wenig* 'a little'. They are part of emerging host-class constructions, consisting of the substantive grammaticalizing elements *ein bisschen*/*ein wenig* and an open slot that can contain nouns, verbs and adjectives. These constructions form the most basic level of description relevant for the grammaticalization of the degree modifiers at hand, and are as such called 'micro-constructions' by Traugott & Trousdale (2013, p. 16). Neels & Hartmann (this volume) argue that both micro-constructions entered into a paradigmatic relation and could be subsumed under a more schematic construction, with an open slot that abstracts over the grammaticalizing elements and an open slot that incorporates the host-class elements. Traugott & Trousdale (2013, p. 14) call such constructions 'schemas' which "abstract across sets of constructions which are (unconsciously) perceived by language-users to be closely related to each other in the constructional network".

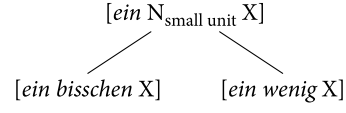


Figure 1. A taxonomic network for German degree modifiers

The above account raises the question how a schema is formed out of lower-level micro-constructions. In general, links in a constructional network are based on similarity, in form and meaning (Diessel, 2015). Neels & Hartmann (this volume) mention some points of similarity between the degree modifier constructions with *ein bisschen* and *ein wenig* that may have spurred the formation of a new schema, i.e. (a) the lexical elements *Bisschen* and *wenig* share the meaning component 'small (unit)', (b) both micro-constructions contain the same substantive element *ein*, and, (c) both constructions have similar highly schematic host-classes.

The schema above includes two open slots, one that abstracts over the host-class slots in the micro-constructions, and one that captures (an embryo of) the paradigm of degree modifiers in German. These slots differ very much in num-

ber of elements, which raises the question whether they differ as to their internal structure. Coussé (2018) has argued that the open class of elements collocating with the grammaticalizing element in host-class constructions is internally structured as a prototype category. Guardamagna (this volume) finds such a prototypical patterning among a group of nouns collocating with the preposition *secundum* 'according to' in Latin. It is an open question whether something similar applies for a paradigm slot with grammaticalizing elements. Kinn (this volume) suggests that there indeed might be a prototypical structure in the first verb slot of the pseudocoordinated construction verb phrase coordination construction [VP_1 *og* VP_2] in Norwegian, with some members more central than others in terms of frequency and signs of grammaticalization. However it may be, one thing both slots seem to have in common is their ability to attract new members.

Bisang (1998, p. 36) argues that open slots in schemas can function as 'attractor positions' that "operate as a kind of melting pot or as a kind of catalyst for linguistic items to be grammaticalized into different types of grammatical functions". Neels & Hartmann (this volume) show that the degree modifier schema serves as a model for new modifiers such as *ein Funken* 'a spark', *eine Idee* 'an idea', *ein Tich* 'a tick', *ein Quänchen* 'a quantum', *eine Preise* 'a pinch' which originally all expressed the meaning of 'small (unit)' and allowed for the indefinite article *ein*. These recent degree modifiers are less frequent and have more restricted collocational patterns than *ein bisschen* and *ein wenig*, which seem to tie in to the findings of Kinn (this volume) for the first verb slot in the asymmetric verb phrase coordination construction [VP_1 *og* VP_2].

The formation of a schema also seems to come with a decrease of formal variation in micro-constructions (Hilpert, 2013). Neels & Hartmann (this volume) show that the use of *ein* with *bisschen* increases in comparison to definite articles, possessives or demonstratives. Enghels & Comer (this volume) similarly find that the use of the preposition following the inchoative auxiliary *poner* gradually becomes fixed to *a*, leading to the loss of the alternate preposition *en* in this position. This tendency seems counterintuitive at first, since the formation of a schema comes with a higher degree in schematicity, i.e. the ability to abstract over a wider set of elements. But on closer look, the reduction of formal variation mostly seems to affect substantive elements, and as such contributes to a 'carving out' of substantive positions in the schema.

5. Construction grammar and its cognitive commitment

Construction grammar has been presented as a framework that allows us to integrate different aspects of grammaticalization. It was shown how the syntagmatic and paradigmatic aspects of grammaticalization may be addressed in a principled way by means of constructions, which are either studied in their own right (syntagmatic axis) or in relation to each other in a taxonomic network (paradigmatic axis). This construction grammar approach also proved capable of handling grammaticalization both as a process of reduction and expansion.

Hilpert (this volume) brings up the issue that construction grammar predominantly serves as a descriptive framework in constructionist approaches to grammaticalization. He points out that construction grammar has a theoretical commitment that goes well beyond this, representing "a mentalistic approach to language that aims to describe speakers' knowledge of language". This cognitive commitment is often left implicit or is altogether missing in constructionist studies of grammaticalization, something that also holds true for the articles in this volume. Hilpert (this volume) therefore raises an open question to diachronic construction grammar: should its central objective be "an explicit commitment to psychological reality" or "to offer a framework for descriptions of language change in which discussions of form and meaning, context, schematicity, and productivity have a natural central place".

A related issue is that many flavors of construction grammar come with a usage-based model of grammatical description (e.g. Langacker, 1987, 2000; Kemmer & Barlow, 2000; Bybee, 2006, 2010). The premise of usage-based theory is that the cognitive representation of grammar is not isolated from its usage (as is the case in generative grammars) but rather is based on and shaped by usage. Von Mengden & Coussé (2014, p. 4–8) point out that a lot of research in grammaticalization is compatible with a usage-based perspective. Types and mechanisms of change such as pragmatic inference, bridging contexts, analogy and frequency effects all explore how usage gives rise to innovative linguistic structures. Following this usage-based tradition, Neels & Hartmann (this volume) investigate the role of bridging contexts for the grammaticalization of degree modifiers. Some of the usage-related concepts from grammaticalization research have also been interpreted in a constructional framework. Pragmatic inference has been analyzed as a 'mismatch' between form and meaning properties leading to a loss of compositionality (Andersson, 2014; Granvik, this volume) whereas semantic bleaching is interpreted as 'coercion by override' resulting in new more procedural or 'bleached' constructions (Booij & Audring, 2018). However, it should be pointed out that not many of these studies take a full-blown usage-based perspective, considering that usage-based grammar implies "grammar based on usage but located

and processed in the human mind" (Von Mengden & Coussé, 2014, p. 8), which is in essence a cognitive perspective.

The descriptive focus of many constructionist approaches to grammaticalization also often implies that these studies do not explicitly commit themselves to one particular flavor of construction grammar. This is also the case in the current volume. One notable exception is Karlsson (this volume) who explicitly chooses Radical Construction Grammar (Croft, 2001) as his theoretical framework. Nevertheless, it is possible to get an insight in the theoretical orientation of most studies in this volume by having a look at their references to standard works and handbooks in construction grammar. The articles in this volume predominantly cite Langacker (1987, 1991), Goldberg (1995, 2006), Croft (2001) and Bybee (2010). These well-known references have in common that they present a usage-based perspective to construction grammar, which has been argued to be compatible with many insights from grammaticalization research. It is remarkable that more formally oriented strands of construction grammar, such as Berkeley Construction Grammar or Sign-Based Construction Grammar (as well as the other formal models presented in Hoffmann & Trousdale, 2013), remain uncited in the articles in this volume, despite their potential for the study of grammatical change (as demonstrated most notably by Fried, 2009, 2013). This might have to do with the general functional orientation of many scholars working on grammaticalization shunning too rigorous formalization.

6. Areas where grammaticalization and construction grammar might not meet

This introduction has focused until now on the advantages of bringing together grammaticalization research and construction grammar, reflecting the general favorable attitude of the articles in this volume. However, there are some voices in the field (especially within grammaticalization research) that offer a critical note, identifying some areas where grammaticalization and construction grammar (might) prove to be incompatible.

Trousdale (2015) and Heine et al. (2016) have identified some theoretical premises of construction grammar that make the framework less suitable for addressing some of the central issues in grammaticalization. Construction grammar, for one, does not make a fundamental distinction between lexical and grammatical forms. This makes the framework incompatible with the very premise of grammaticalization research which addresses "questions as how lexical items and constructions come in certain contexts to serve grammatical functions or how grammatical items develop new grammatical functions"

(Hopper & Traugott, 2003, p. 1). A related issue is that not all constructionist approaches consider morphemes to be constructions, i.e. independent pairings of meaning and form (e.g. Booij, 2010, p. 15). This feature seriously hampers the integration of the concept of grammaticalization clines, e.g. the change from content item → grammatical word → clitic → inflectional affix (Hopper & Traugott, 2003, p. 7), into a construction grammar perspective.

Heine et al. (2016, p. 160) also point out some "issues that are important to students of grammaticalization while they have less or no significance in work based on Construction Grammar". As the discussion below will show, this claim is too strong, and erroneously narrows down the scope of construction grammar to the study of constructions in isolation. Heine et al. (2016, p. 150) for instance make a case for 'event schemas' as the cognitive-conceptual foundation of grammatical categories, claiming that these schemas differ "from the symbolic form-meaning units of Construction Grammar […] in the fact that they are defined as cognitive-conceptual categories". True, event schemas are not the same as constructions, but they do play an important role in construction grammar. First, the meaning of constructions may be described in terms of event schemas, as is demonstrated by the cognitive schemas for posture verbs in Lesuisse & Lemmens (this volume). Second, the meaning of individual constructions may be mapped onto a more abstract conceptual space, as has been shown in great detail by Croft (2002).

Heine et al. (2016) go on to argue that semantic changes in grammaticalization are driven by discourse and context rather than the construction itself. Again, this issue is not incompatible with construction grammar, rather on the contrary. It was argued in the preceding section that many strands of construction grammar take a usage-based perspective on grammar, implying that constructions are based on and shaped by usage (including discourse and context). Neels & Hartmann (this volume) illustrate this usage-based approach by exploring the role of bridging contexts for the grammaticalization of degree modifiers. Finally, Heine et al. (2016) claim that the "main arena of work on constructional change is language history, that of students of grammaticalization is the search for typological generalizations on the evolution of functional categories". This assessment contrasts with the effort of Hölzl (this volume) to translate Heine & Kuteva's (2010) concept of 'grammaticalization areas' into 'constructionalization areas', i.e. "a group of often geographically contiguous languages that share an analogous instance of constructionalization due to language contact".

7. Overview of the contributions

Now that the theoretical groundwork of this volume is laid, it is time to present the individual articles in more detail. The volume opens with a theoretical contribution by **Martin Hilpert** raising three open questions to the diachronic study of constructions and constructional change: (1) To what degree should practitioners of diachronic construction grammar make psychological claims when using constructions as the basic unit of their analysis? (2) When does an emerging construction count as a new construction? (3) Is there such thing as form-meaning nodes in the constructional network?

These three questions are of great interest and their potential answers have a great impact on basic concepts of grammatical theory and how to understand change in the grammatical system. Hilpert argues that aspects of diachronic construction grammar need to be clarified explicitly. For one, there is need for some meta-theoretical reflection on how diachronic construction grammar can make claims about the linguistic knowledge of earlier generations.

Furthermore, the status of a new construction is problematic as it can only be decided in relation to a starting point that is chosen by the analyst. An alternative way to distinguish different types of change could be to ask which parts of the network are affected by a given process of change. As to the third question, Hilpert points out some directions for future research, calling for a shift of perspective. He argues for a connection-centered view of linguistic knowledge, which reduces constructions to being links between form and meanings, rather than being nodes in the network.

The remaining contributions in the volume have a more empirical focus, each of them tackling one particular case of grammaticalization or constructional change in one particular language or language family. These articles are arranged thematically around the type of construction in focus.

7.1 Verb constructions

Mégane Lesuisse and **Maarten Lemmens** investigate the historical development of the posture verbs *sit*, *stand* and *lie* in English. This trio of verbs is fascinating as they have not (or only very little) grammaticalized into locative markers or copulas, which sets them apart from their cognates in other Germanic languages. Lesuisse & Lemmens search for traces of grammaticalization of these posture verbs in historical dictionaries and corpora. They combine qualitative methods with quantitative productivity measures in order to uncover signs of expansion in the locative and copular use of the three verbs. The data analysis reveals that posture verbs did occur in locative and copula constructions in the history of English,

which indicates that their "grammaticalization was cut short" at some point in history. Lesuisse & Lemmens try to explain this halted grammaticalization by testing the hypothesis of Newman (2009) that it is the innovative use of posture verbs expressing dynamic events (getting into a sitting, standing, lying posture) that has hampered their further grammaticalization.

Torodd Kinn addresses an interesting family of verb constructions in Norwegian, pseudocoordinated constructions [VP_1 *og* VP_2], where the first verb in some cases may express grammatical meaning. The construction is particularly interesting since it has multiple meanings and shows great variation in the verb slots VP_1 and VP_2. The paper focuses on the properties of a set of subschemas expressing atelic motion as in *løpe rundt og leke* 'run around and play'. An extensive corpus study in Present-day Norwegian shows (in contrast to earlier thinking) that a great range of verbs may fill the VP_1 slot; some verbs however are more central than others. Using quantitative methods from construction grammar, such as distinctive collexeme analysis, Kinn shows that the bleaching of particular verbs in VP_1 correlates with non-agentive (mainly cognitive) verbs in VP_2. He argues for a complex interplay between bleaching of particular verbs and constructional semantics, all correlating with decreased argument structure inside the construction. As a result of the semantic bleaching of some verbs in VP_1, the facilitation relation that holds between VP_1 and VP_2 is blurred, which also affects the meaning of the construction as a whole, since VP_2 contributes most of the lexical meaning, and this strengthens the single-event meaning that already characterizes pseudocoordination in general.

Renata Enghels and **Marie Comer** examine the grammaticalization of the Spanish verbs *poner* and *meter* 'put' into inchoative auxiliaries. Inspired by Hilpert (2013), they explore the structural variants of the inchoative construction [NP_1 V_{refl} PREP INF] in order to examine "to what extent different structural patterns have given rise to a general constructional schema". To this purpose, they perform a corpus study taking into account (a) the general productivity of the construction, (b) the formal variation in the prepositional slot PREP, (c) the occurrence of intervening adverbials in the construction, and, (d) the overall semantics of the construction. Their data analysis shows that the inchoative construction in Spanish becomes more productive, shows less variation in the prepositional slot, allows less intervening adverbials, and is used in a wider array of contexts. Enghels & Comer take these findings as evidence "that a leveling process has taken place, and that an abstract constructional inchoative schema has been formed". They also point out that this process is more advanced in constructions with *poner* in comparison to the semantically specialized verb *meter*.

7.2 Noun constructions

Jakob Neels and **Stefan Hartmann** address the grammaticalization of the degree modifiers *ein Bisschen* 'a bit' and *ein wenig* 'a little' in German. The history of degree modifiers has been studied from different angles in constructionist approaches, mostly based on English data but also on Dutch and Spanish. Neels & Hartmann build on this research tradition, exploring both syntagmatic and paradigmatic aspects of the grammaticalization of the degree modifiers in German. They first focus on the syntagmatic properties of the host-class constructions [*ein bisschen* X] and [*ein wenig* X], investigating (a) the general productivity of the constructions, (b) the parts of speech found in the host-class X, and, (c) the formal variation of the determiner *ein* in [*ein bisschen* X]. Their corpus results show a "reduction in individual formal variation and expansion in their range of use" as such uncovering both reduction and expansion aspects of grammaticalization at work simultaneously. The micro-constructions [*ein bisschen* X] and [*ein wenig* X] moreover show an increasing convergence, both functionally and formally, which leads Neels & Hartmann to posit the formation of a constructional schema [*ein* $N_{small\ unit}$ X]. This schema is hypothesized to function as a model for new degree modifiers in German.

Caterina Guardamagna takes a constructionist approach to diachronic changes in the Latin [*secundum* NP] 'according to NP' construction. She elaborates on type frequency effects such as productivity and schematicity (Barðdal, 2008), as well as host-class expansion (Himmelmann, 2004). Her investigation is based on empirical data from the Latin Library Corpus, covering data from four periods: Classical Latin (BC106–AD17), Silver Latin (100–258), Late Latin (330–469) and Early Medieval Latin (530–704). The analyzed data, consisting of nearly 1000 tokens, shows that the [*secundum* NP] construction becomes more productive over time. Guardamagna argues that constructional change includes frequency changes. As for the host-class expansion, she argues that social context stimulates some of the changes, in the sense that the [*secundum* NP] construction with a 'conformity' meaning connects to cultural changes in the transition from Paganism to Christianity.

7.3 Constructions at sentence level

Anton Granvik investigates the development of the abstract noun *caso* 'case' into a conditional marker in Spanish, e.g. *(en) (el) caso (de) que llueva, no saldremos* '(in) (the) case (of) that it rains, we'll stay in'. In accordance with the theme of this volume, he compares how grammaticalization and constructionalization models fare in accounting for the creation of a new grammatical meaning and

the subsequent formal changes in the conditional *caso* construction. In his discussion, he finds that both models are accurate, but serve for different aspects of the development of the construction. As for the development of the conditional reading of *caso*, the grammaticalization model provides an accurate explanation for describing the directional change. However, concerning formal changes in the construction, Granvik finds the constructionalization and constructional change model of Traugott & Trousdale (2013) more suitable. Especially, the network model of construction grammar provides a powerful tool in this case, since it allows the researcher to relate different forms to higher and lower level schemas.

Andreas Hölzl starts out from the concept of 'grammaticalization areas' defined by Heine & Kuteva (2010, p. 97) as "a group of geographically contiguous languages that have undergone the same grammaticalization process as a result of language contact". Hölzl argues for the existence of 'constructionalization areas' as well, that is similar changes in constructions and constructional networks among contiguous languages due to contact. He examines the concept of negation in Tungusic and Mongolian languages in Northeast Asia, focusing on Manchu and Mongolian. By examining a great range of constructions and constructs with similar subsequent constructional changes, Hölzl shows that the languages addressed in share a development from negative existential (non-existence of x) to a standard negator and as such a development towards a more abstract and procedural meaning. Hölzl emphasizes the importance of a shared source construction, arguing that it is the construction as a whole that expands its domain of use rather than the negative existential alone. He subsumes all changes under shared constructionalization and discusses the development in terms of an "extension of the interlingual constructional network" shared by the languages included.

Emanuel Karlsson makes an intriguing attempt to combine insights from radical Construction Grammar (Croft, 2001) with grammaticalization, investigating the diachrony of spatial particles from Proto-Indo-European to Ancient and Classical Greek, such as *epi* 'on' and *kata* 'down'. The empirical data show a semantic shift of *epi* and *kata* across constructions, and the particles develop from being used more or less as spatial adpositions to verbal prefixes (expressing aspect) and independent adverbs. Karlsson uses the term 'construction split' for the development of multiple constructions and argues for a 'pragmatics-driven' gradual development where there is no single point of reanalysis of covert structure. In taking a pragmaticist position, he emphasizes the importance of concepts such as 'utterance cues' (rather than form or substance) and 'discourse effects' (rather than function or meaning). Karlsson suggests that the senses of individual particles develop in different directions under the influence of the specific constructions in which they are used, and that the fundamental gradualness of synchrony and diachrony correspond to one another. As such, the paper attempts a partial

synthesis of theoretical components relevant to diachronic construction grammar and grammaticalization from the perspective of a cue-based approach to constructions.

References

Andersson, P. (2014). The fast case. Constructionalization of a Swedish concessive. *Nordic Journal of Lingustics*, 37, 141–167. https://doi.org/10.1017/S033258651400016X

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij, & J. van Marle (Eds.), *Yearbook of morphology 1992* (pp. 181–208). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-3710-4_7

Baayen, H., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, 29, 801–844. https://doi.org/10.1515/ling.1991.29.5.801

Barðdal, J. (2008). *Productivity. Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8

Barðdal, J., Smirnova, E., Sommerer, L., & Gildea, S. (Eds.). (2015). *Diachronic Construction Grammar*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.18

Bergs, A., & Diewald, G. (Eds.). (2008). *Constructions and Language Change*. Berlin: de Gruyter. https://doi.org/10.1515/9783110211757

Bisang, W. (1998). Grammaticalization and language contact, constructions and positions. In A. Giacalone Ramat, & P. J. Hopper (Eds.), *The Limits of Grammaticalization* (pp. 13–58). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.37.02bis

Booij, G. (2010). *Construction morphology*. Oxford: Oxford University Press.

Booij, G., & Audring, J. (2018). Category change in Construction Morphology. In K. Van Goethem, N. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category change from a constructional perspective*. Amsterdam: John Benjamins.

Bybee, J. (2003). Mechanisms of change in grammaticization. The role of frequency. In R. Janda, & B. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell Publishers. https://doi.org/10.1002/9780470756393.ch19

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

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Coussé, E. (2018). Grammaticalization, host-class expansion and category change. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), Category change from a constructional perspective. Amsterdam: John Benjamins.

Croft, W. (2001). *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Croft, W. (2002). *Typology and Universals*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511840579

Diessel, H. (2015). Usage-based construction grammar. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110292022-015

Diewald, G. (2002). A model for relevant types of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 103–120). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.09die

Enghels, R & Comer, M. (this volume). Evaluating grammaticalization and constructional accounts: The development of the inchoative construction with put verbs in Spanish

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions. The case of *Let alone* . *Language*, 64, 501–538. https://doi.org/10.2307/414531

Fried, M. (2009). Construction Grammar as a tool for diachronic analysis: Constructions and Frames 1(2), 261-290.

Fried, M. (2013). Principles of Constructional Change. In: G Hoffmann & G. Trousdale (2013). *The Oxford Handbook of Construction Grammar*, (pp. 1–21). Oxford: Oxford University Press.

Gisborne, N., & Patten, A. (2011). Grammaticalization and Construction Grammar. In H. Narrog, & B. Heine (Eds.), *The Oxford Handbook of Grammaticalization* (pp. 92–104). Oxford: Oxford University Press.

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

Goldberg, A. (2006). *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Granvik, A. (this volume). The development of the conditional caso construction i spanish.

Guardamagna, C. (this volume). Type frequency, productivity and schematicity in the evolution of the Latin secundum NP construction

Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 83–101). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Heine, B., & Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (Ed.), *Handbook of language contact* (pp. 86–105). Malden: Wiley-Blackwell. https://doi.org/10.1002/9781444318159.ch4

Heine, B., Narrog, H., & Long, H. (2016). Constructional change vs. grammaticalization. From compounding to derivation. *Studies in Language*, 40, 137–175.

Hilpert, M. (this volume). Three open questions in diachronic construction grammar.

Hilpert, M. (2013). *Constructional changes in English. Developments in allomorphy, word formation and syntax*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139004206

Himmelmann, N. P. (2004). Lexicalization and grammaticalization. Opposite or orthogonal? In W. Bisang, N. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization? A Look from its Fringes and its Components* (pp. 21–42). Berlin: Mouton de Gruyter.

Hoffmann, Th., & Trousdale, G. (2013). *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195396683.001.0001

Hopper, P., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Hölzl, A. (this volume). Constructionalization areas: The case of negation in Manchu.

Israel, M. (1996). The way constructions grow. In A. Goldberg (Ed.), *Conceptual Structure, Discourse and Language* (pp. 217–230). Stanford: CSLI Publications.

Karlsson, E. (this volume). A Radical Construction Grammar approach to construction split in the diachrony of the spatial particles of Ancient Greek: Some theoretical preliminaries.

Kemmer, S., & Barlow M. (Eds.). (2000). *Usage-based models of language*. Stanford: CLSI Publications.

Kinn, T. (this volume). Pseudocoordination in Norwegian: Degrees of grammaticalization and constructional variants

Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume I. Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, R. W. (1991). *Foundations of cognitive grammar. Volume II. Descriptive application*. Stanford: Stanford University Press.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). Stanford: CLSI Publications.

Lehmann, C. (1982). *Thoughts on Grammaticalization. A programmatic Sketch*. Volume I. Arbeiten des Kölner Universalien-Projekts 48. Köln.

Lesuisse, M & Lemmens, M. (this volume). Grammaticalisation cut short: A diachronic constructional view on English posture verbs

Newman, J. (2009). English posture verbs: an experientially grounded approach. *Annual Review of Cognitive Linguistics*, 7, 30–57. https://doi.org/10.1075/arcl.7.02new

Neels, J & Hartmann, S. (this volume). Reduction or expansion? A bit of both: A case study on the development of German degree modifiers.

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

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional change*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Traugott, E. C. (2003). Constructions in grammaticalization. In B. D. Joseph, & R. D. Janda (Eds.), *The Handbook of Historical Linguistics* (pp. 624–647). Oxford: Blackwell. https://doi.org/10.1002/9780470756393.ch20

Trousdale, G. (2015). *Three areas where Grammaticalization and Construction Grammar don't meet*. Paper presented at the workshop 'Grammaticalization meets Construction Grammar', Gothenburg 8–9 October 2015.

Tuggy, D. (2007). Schematicity. In D. Geeraerts, & H. Cuyckens (Eds.), *Oxford Handbook of Cognitive Linguistics* (pp. 82–166). Oxford: Oxford University Press.

Von Mengden, F., & Coussé, E. (2014). The role of change in usage-based conceptions of language. In E. Coussé, & F. von Mengden (Eds.), *Usage-based approaches to language change* (pp. 1–19). Amsterdam: John Benjamins.

Three open questions in Diachronic Construction Grammar

Martin Hilpert
University of Neuchâtel

A growing number of studies on language change adopt Construction Grammar as a theoretical framework so that there is now a developing field of Diachronic Construction Grammar. As is typical of any emerging linguistic theory, many aspects of Diachronic Construction Grammar are still not clarified explicitly, or they are understood in different ways by different researchers. There is thus a need for some meta-theoretical reflection. This paper identifies three questions in Diachronic Construction Grammar that are currently unresolved. These three questions pertain to the status of constructions as mental representations of language, the emergence of new constructions, and the way in which nodes and connections are viewed as parts of the constructional network. By airing these questions, this paper hopes to stimulate further discussion.

1. Introduction

Over the past decade, Construction Grammar (Goldberg, 1995, 2006) has become an increasingly popular theoretical framework for the analysis of language change (see for instance Noël, 2007; Bergs & Diewald, 2008; Traugott & Trousdale, 2013; De Smet, 2013; Hilpert, 2013; Petré, 2014; Barðdal et al., 2015; Heine et al., 2016; amongst many others). While Diachronic Construction Grammar is undoubtedly undergoing a very promising development, many of its aspects are still not clarified explicitly, or are understood in different ways by different researchers. There is thus a need for some meta-theoretical reflection, to which this paper aims to contribute. Specifically, this paper points to three questions that are currently unresolved and that deserve the attention of researchers who are interested in the further development of Diachronic Construction Grammar.

The first question that will be raised concerns the status of constructions as mental representations of linguistic structure. Construction Grammar, as a synchronic theory of language, has the goal of describing speakers' knowledge of

https://doi.org/10.1075/cal.21.c2

language, and there have been notable advances in usage-based constructional approaches that have linked frequencies from corpus data to the notion of entrenchment and other aspects of linguistic knowledge (Bybee, 2010). Despite these advances, it is clear that historical corpora, as the main source of data for diachronic analyses, give us only a very rough idea of language use in the past. It is therefore an open question how confidently we can make statements about the linguistic knowledge of earlier generations of speakers, and whether this is actually the main goal of Diachronic Construction Grammar.

A second pertinent issue is the phenomenon that Traugott & Trousdale (2013) call constructionalization, i.e. the creation of a new node in the speaker's mental network of constructions. With historical corpus data, it is of course possible to detect novelties in language use that are inconsistent with prior regularities and constraints on the use of a particular linguistic form. In this way, approximate dates can be determined for the emergence and spread of new constructions. However, the concept of constructionalization is also open to criticism. Traugott & Trousdale distinguish between constructionalization, which is the creation of a new node, and constructional change, which is a change that affects an already existing node. This differentiation raises the question just after how many constructional changes exactly we have a construction that counts as a new node. The term, as defined, asks us to think of a discrete threshold. Whether such a threshold exists is an open question.

The third question that this paper addresses does not just concern Diachronic Construction Grammar, but in fact the field of Construction Grammar as a whole. It appears to be a largely unquestioned consensus across constructional and usage-based approaches that linguistic knowledge is to be modeled as an associative network in which there are nodes, i.e. constructions, and links between those constructions. Schmid (2017) argues for a view in which knowledge of language exclusively takes the format of associations, which effectively reduces constructions to links between forms and meanings, relegating the nodes to a rather marginal status. This proposal is not primarily motivated by theoretical parsimony, but rather by the aim of describing linguistic knowledge in inherently dynamic terms. It is clear that this idea has profound implications for the constructional study of language change, some of which this paper will explore.

Each of the three questions will be dealt with in a separate section below. A final section will identify issues where the questions intersect, draw together the most important ideas, and formulate some conclusions. It is clear that this paper will not deliver final answers to these questions. It might even turn out that they are the wrong questions to ask in the first place. Hopefully though, this paper will contribute to the ongoing discussion about where the field is going, thereby stimulating new and exciting research in Diachronic Construction Grammar.

2. What is investigated in Diachronic Construction Grammar?

Construction Grammar represents a mentalistic approach to language that aims to describe speakers' knowledge of language (Hilpert, 2014, Chapter 1). The objects of analysis are typically one or more constructions, understood as generalizations, that is, speakers' mental representations of how linguistic forms are coupled with meaning. Constructional analyses may investigate what idiosyncrasies and constraints motivate the recognition of a form-specific generalization such as the *let alone* construction (Fillmore et al., 1988), or they may seek to demonstrate the psychological reality of generalizations across several linguistic forms and their meanings, as in the case of the ditransitive construction and the prepositional dative construction (Perek, 2015). In both scenarios, the argument for the existence of a given generalization has to be made on the basis of behavioral evidence. Either there is behavioral data from an experiment, indicating that speakers act in accordance with a constructional generalization, or there is corpus-linguistic data that, in line with the tenets of usage-base linguistics (Bybee, 2010), supports the idea that speakers' language use reflects a constructional generalization. If Diachronic Construction Grammar were simply regular Construction Grammar with the inclusion of change over time as an additional dimension, then its main goal should be to investigate how knowledge of language changes over time. Yet, this perspective is not particularly dominant in many studies that address the diachrony of constructions, and it could even be argued that any claims about the linguistic knowledge of earlier generations of speakers stands on rather shaky ground, given the limited representativeness of historical corpora.

Noël (2007, p. 178) explains the popularity of Diachronic Construction Grammar with reference to the fact that "many functionalists and cognitivists have long been working with a pre-theoretical constructional notion", meaning that the term construction can be conveniently used in order to talk about linguistic forms, their meanings, and their historical developments, even when theory-specific considerations such as speakers' mental representations of language are not at stake and no claim about these representations is being made. Also von Mengden & Coussé (2014, p. 9) point out that adopting a cognition-centered perspective is not a necessary precondition for diachronic studies in usage-based linguistics. In other words, Diachronic Construction Grammar can be chosen as a descriptive framework that favors a perspective on changes in individual form-meaning pairs, but any reference to the psychological reality of these form-meaning pairs can be left implicit. Likewise, the descriptive nature of the constructional approach leaves open to what extent a researcher is committed to hypotheses and explanatory principles that are held in frameworks such as grammaticalization theory (Hopper & Traugott, 2003), which make strong empirical predictions with

regard to language change. This kind of flexibility does of course render Construction Grammar vulnerable to the criticism that an absence of testable predictions makes it hard for others to falsify aspects of the model. In practice, however, it appears that this drawback has not diminished the popularity of the framework.

Another feature that contributes to the current popularity of the constructional framework is that it allows analysts to integrate descriptions of formally specific constructions with analyses of more abstract patterns. In the words of Traugott (2015, p. 52), "the architecture of construction grammar demands thinking in terms of both meaning and form, and not only of individual substantive constructions but also abstract schematic ones". While the notion of a schematic pattern strongly invites the idea of a speaker who mentally represents an abstraction over many concrete usage events, it is actually possible for an analyst to remain agnostic about this issue. It can be left open whether abstract patterns have psychological reality or whether they are in fact merely posited as theoretical constructs, much in the way that linguistic notions such as 'head', 'clitic', or 'subordinating conjunction' serve as useful labels for phenomena that capture insights about a linguistic system but that may not correspond to psychologically real categories in the minds of speakers.

While it is certainly non-trivial to use historical corpus data for the purpose of making inferences about the linguistic knowledge that speakers of earlier historical periods may have had, there are studies that offer promising results. For example, Wolk et al. (2013) adapt quantitative methods from variationist linguistics in order to show that many of the factors that guide present-day speakers' choices between the two constructions of the dative alternation were also at work in Late Modern English. Similarly, Sommerer (2015) draws on data from the Peterborough Chronicle and the Parker Chronicle to make the case that speakers started to represent an abstract NP construction with an obligatory slot for a determiner, which became increasingly strongly entrenched. In my own work I have used distributional evidence to suggest that parenthetical concessive clauses in recent American English (as in *This is an interesting, if problematic, suggestion*) are not mentally represented as a single, general construction, but rather as a family of constructions at lower levels of abstraction (Hilpert, 2013, p. 199). What these examples show is that Diachronic Construction Grammar with a cognitive outlook is possible in principle, even though there are of course many restrictions. What is less clear is whether this perspective actually captures the main goal of a constructional approach to language change.

In this context, it is important to keep in mind that many researchers in Diachronic Construction Grammar come from a background of grammaticalization theory, which is a research program that of course has strong connections to cognitive concerns (Heine, 1997; Bybee, 2003; van der Auwera et al., 2015;

Hilpert, 2015a 2015b), but which is just as strongly concerned with structural descriptions of language at a level that goes well beyond the confines of a single human mind. Issues such as the hypothesis of unidirectionality (Janda, 2001), the existence of cross-linguistic clines of diachronic change (Bybee et al., 1994), or the development of paradigms (Nørgård-Sørensen & Heltoft, 2015) can be fruitfully investigated without placing cognition at the center of the discussion. The 'constructional turn' in grammaticalization studies, labeled in this way by van Rompaey et al. (2015), arguably does not reflect an increased interest in the cognitive ramifications of grammaticalization, but rather a greater awareness of the fact that syntagmatic relations play an important role in grammatical change. The notion of a construction, which Construction Grammar provides, thus gives greater precision to descriptions of grammaticalizing forms, allowing for the distinction of different morphosyntactic contexts in which grammaticalization does or does not take place, whereas the implications that the term has with regard to cognition often remain unaddressed.

The open question that I would like to raise in this section is whether there ought to be an explicit commitment to psychological reality, as for instance in the 'cognitive commitment' that is upheld in Cognitive Linguistics (Evans, 2016), or whether the first priority of Diachronic Construction Grammar should simply be to offer a framework for descriptions of language change in which discussions of form and meaning, context, schematicity, and productivity have a natural central place. While these two aspects are not necessarily in mutual conflict, it should be clear that any claims about cognition will incur the cost of having to make them plausible with evidence from historical corpora or from experiments with present-day speakers. While greater restrictiveness is of course a desirable quality of scientific theories, imposing a strict cognitive commitment on Diachronic Construction Grammar might also have a detrimental effect on the field. There are researchers who appreciate the descriptive tools of Construction Grammar but whose studies are primarily directed at language structure, not cognition. It would be a severe mistake to ignore the insights that are provided by studies of this kind. By contrast, it would probably be useful if researchers with a commitment to investigating cognition were to state this goal explicitly. Not only would this lead to enhanced transparency, but it would also give others a chance to come up with contradicting evidence, thus improving the falsifiability of diachronic constructionist studies.

3. When is a new construction a new construction?

One of the central theoretical claims of Construction Grammar is that speakers' knowledge of language is organized in a large network in which form-meaning pairs of varying complexity and schematicity are hierarchically interconnected (Goldberg, 2003, p. 219). This idea is widely shared in Diachronic Construction Grammar, which raises the question of how the constructional network can change over time. Language change, as viewed from the perspective of Construction Grammar, would instantiate one of several types of change in the constructional network. This section will be primarily concerned with the emergence of new constructions in the network, which is a process that Traugott & Trousdale (2013) call constructionalization. Not only can new constructions appear in the network, but of course, old constructions may also disappear. Another dynamic process that may affect nodes in the network has been called constructional change, which Hilpert (2013, p. 16) defines as a change that "selectively seizes a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these". Crucially, in comparison to Traugott & Trousdale's constructionalization, the creation of new nodes is not a necessary result of constructional change. Changes that have, at least up to now, not received a great deal of attention are changes that affect connections in the network, rather than nodes (but see the discussion of connection-centered work at the end of ection 4 in this paper). Over time, new connections can appear, for instance as existing forms become conventionally associated with new meanings. For example, many present-day speakers of English have come to associate the form *tablet* with the meaning 'portable computer without a keyboard'. By the same token, connections that used to exist at some point may disappear from the network (*meat* can no longer refer to 'food in general'). Both processes involve the gradual strengthening or weakening of connective links, so that for instance a polysemous construction may change over time with regard to its connections between its form and its meanings and their respective strengths (*fantastic* nowadays has a stronger connection to 'wonderful' than to 'imaginary'). While a thorough systematization of all of these types of change still remains to be worked out, Traugott & Trousdale (2013, p. 22) provide a detailed definition of the change that is perhaps the most interesting and relevant, namely the emergence of new constructions:

> Constructionalization is the creation of $form_{new}$-$meaning_{new}$ (combinations of) signs. It forms new type nodes, which have new syntax or morphology and new coded meaning, in the linguistic network of a population of speakers. It is accompanied by changes in degree of schematicity, productivity, and compositionality.

> The constructionalization of schemas always results from a succession of micro-steps and is therefore gradual. New micro-constructions may likewise be created gradually, but they may also be instantaneous. Gradually created micro-constructions tend to be procedural, and instantaneously created micro-constructions tend to be contentful.

The terms *procedural* and *contentful* capture a distinction between constructions that could be called grammatical and lexical, respectively. The definition states that grammaticalization, or grammatical constructionalization in the terminology of Traugott & Trousdale, is a gradual process, whereas the creation of new lexical items can be instantaneous. An important addendum to the definition is the statement that "[f]ormal changes alone, and meaning changes alone cannot constitute constructionalization" (Traugott & Trousdale, 2013, p. 22). Formal change or meaning change in isolation would, according to Traugott & Trousdale, instantiate constructional change – change within an existing node of the network. The concept of constructionalization as the emergence of a new form-meaning pair is intuitively appealing, as there is no shortage of clear examples. New words such as *selfie* or *photobomb* enter the lexicon, new constructions like the *get*-passive come into existence. Traugott & Trousdale make clear that they view examples of the latter kind, i.e. grammatical constructionalization, as a gradual process, which is very much in line with current views in grammaticalization studies and usage-based linguistics. Why then would anyone take issue with the idea of constructionalization?

In the following paragraphs, I will argue that despite Traugott & Trousdale's commitment to gradual change, the term constructionalization ultimately invites the notion of a discrete threshold between an existing construction that has changed and a new construction that has come into being. This threshold may turn out to be a line in the sand that is impossible to draw with certainty. To illustrate the problem with an example, we can consider the recent history of the English verb *confirm*, which has undergone both semantic and syntactic change over the past 150 years, and which hence would instantiate constructionalization, more specifically, lexical constructionalization. Syntactically, it changes its behavior from a verb that has a near-categorical preference for nominal direct objects into a verb that regularly occurs with *that*-clauses (Hilpert, 2011, p. 456). The examples below, taken from COHA (Davies, 2010), show that these different syntactic variants of *confirm* co-vary with different meanings.

(1) This circumstance tended to confirm my fears that Constantia was dead.

(2) Michael? Penzley here. Just calling to confirm that you've got the new Bundy ads ready for us today.

In the first example, the verb *confirm* is used in its sense of 'corroborate, affirm', and the proposition that is affirmed is expressed in the following noun phrase. In the second example, the meaning has been extended metonymically and is hence slightly different, so that *confirm* means 'ask someone else to verify', and the information to be verified is expressed by a *that*-clause. To make the semantic difference clear, the second *confirm* can be paraphrased by *make sure* (*Just calling to make sure that you've got…*), whereas this is not possible in the first example (*?tended to make sure my fears*). The new sense of *confirm* has a clear preference for the more complex syntactic form, so that sentences with a direct object, as for example *I'd like to confirm our reservation*, need heavy contextual scaffolding in order to convey the 'make sure' sense, rather than the 'affirm' sense.

The details of the semantic and syntactic analysis aside, it appears that in Present-Day English, there is a new form-meaning pair that would qualify as a case of lexical constructionalization. What is less clear is whether it is possible to point to a specific set of examples in order to show when exactly the line between constructional change and constructionalization was crossed. The verb *confirm* shows a gradual diachronic increase in its preference for clausal complements, so that in all likelihood, the syntactic change that allowed *confirm* to take *that*-clauses happened long before the new 'make sure' sense emerged. Now, if the semantic change took place during a time when the syntactic change was already complete, this would suggest that we are in fact dealing with isolated cases of formal change and meaning change that do not, after all, instantiate constructionalization. It is unclear whether the two have to occur simultaneously, and if so, what would make simultaneous change of form and function different from a change in which the two happen in sequence. The results, it would seem, are indistinguishable. In both cases, new form-meaning pairs come into being.

In their review of Traugott & Trousdale (2013), Börjars et al. (2015) note exactly this point and pose the question how the emergence of a new form-meaning pair could be analyzed into steps of constructionalization and constructional change. They discuss the idea that constructionalization could be thought of as the final step that creates a form-meaning pair in which both form and meaning have undergone a change. If we use the case of *confirm* as an example, the component changes of its development could be schematized as shown in (3) below. In the example, the double arrow '>>' stands for mere continuation of a form-meaning pair, '>cc>' stands for constructional change, and '>czn>' stands for constructionalization.

(3)

$FORM_1$ – $MEANING_1$	>>	$FORM_1$ – $MEANING_1$	>>	$FORM_1$ – $MEANING_1$		
	>cc>	$FORM_2$ – $MEANING_1$	>>	$FORM_2$ – $MEANING_1$		
			>czn>	$FORM_2$ – $MEANING_2$		

It is only in the second step that a completely new form-meaning pair arises. Importantly, Börjars et al. (2015) point out that the moment of constructionalization is actually dependent on the starting point that the analyst chooses for the development. If an earlier starting point is introduced, say, with a pairing of $FORM_1$ and $MEANING_0$, a completely new form-meaning pair is attested already with the emergence of $FORM_2$ – $MEANING_1$, as is shown in (4).

(4)

$FORM_1$ – $MEANING_0$	>>	$FORM_1$ – $MEANING_0$	>>	$FORM_1$ – $MEANING_0$	>>	$FORM_1$ – $MEANING_0$	
	>cc>	$FORM_1$ – $MEANING_1$	>>	$FORM_1$ – $MEANING_1$	>>	$FORM_1$ – $MEANING_1$	
			>czn>	$FORM_2$ – $MEANING_1$	>>	$FORM_2$ – $MEANING_1$	
					>czn>	$FORM_2$ – $MEANING_2$	

This means that constructionalization is actually a relative term. A completely new form-meaning pairing can only ever be identified relative to a starting point. For *confirm* in the sense of 'make sure', i.e. $FORM_2$ – $MEANING_2$ in (3) and (4) above, this means that it would be the result of constructionalization relative to the first step in (3), but, paradoxically, only the result of constructional change if the second step is adopted as the starting point. The difference between constructionalization and constructional change is thus not a matter of different linguistic processes, but rather a distinction that lies in the eye of the beholder.

The term constructionalization is of course useful as a general label for the emergence of a new construction, where the analyst wishes to highlight the fact that both form and meaning are different from an earlier construction that is chosen as a reference point. Yet, there are probably distinctions that are both less difficult to establish and more useful from a theoretical standpoint than the difference between constructionalization and constructional change. One such distinction relates to the question of how different parts of the constructional network are affected by a given process of change. If we start from the broad consensus that speakers' knowledge of language essentially includes forms, meanings, and connections (between forms and meanings, but also between two or more forms, or

two or more meanings), and if we accept the idea that new parts of the network may emerge, become stronger, become weaker again, and eventually disappear, we can draw up a table in which all of these categories intersect. Table 1 below gives concrete examples for the types of change that are differentiated in such a cross-tabulation.

Table 1. Types of change in the constructional network

	Form	**Meaning**	**Connection**
emergence	new forms appear: *selfie*, *contrafibularities*, the *get*-passive	new meanings /concepts appear: 'wireless internet access', 'gluten intolerance', 'passive-aggressive'	new connections are formed: The existing form *gay* forms a new connection with the existing meaning 'homosexual'.
strengthening	forms gain in strength of representation: (semasiological) frequency increase of *like* as a discourse marker	meanings /concepts gain in strength of representation: (onomasiological) frequency increase of the meaning 'wireless internet access'	connections gain in strength: The form *fantastic* becomes more strongly associated with the meaning 'wonderful'.
weakening	forms decrease in strength of representation: (semasiological) frequency decrease of *whom* as a relative pronoun	meanings /concepts decrease in strength of representation: (onomasiological) frequency decrease of the meaning 'person granted the use of land in exchange for goods and service'	connections fade in strength: The verb *dislike* becomes less strongly associated with the complementation pattern of a following *to*-infinitive.
disappearance	forms disappear: The form *affuage* is no longer used.	meanings /concepts disappear: The meaning 'the right to cut firewood in a forest' is no longer expressed with a single word such as *affuage*.	connections disappear: The English ditransitive construction is no longer associated with the meaning of banishment or exclusion (e.g. *He therefore forbade her the court*).

The first two columns of the table, which pertain to form and meaning, could be grouped together under the heading of 'node changes'. Importantly, the separate presentation of formal changes and meaning changes in the table does not preclude developments in which form and meaning change in conjunction with each other. Simultaneous changes of this kind are in fact inevitable when new words and constructions enter a language, or when linguistic elements fall out of

use. The third column describes processes that instantiate 're-wiring changes' in the constructional network. Clearly, the examples in Table 1 do not exhaust all possible types of change that could be observed in the constructional network, but they illustrate the general logic of identifying the respective part of the network that is affected and specifying how that part changes over time. The right-hand column holds particular promise for a further development of the model. Many interesting processes that relate to terms such as grammaticalization and constructionalization could be re-cast as connectivity changes in which the network undergoes re-wiring. For example, the phenomena of increasing schematicity and productivity (cf. Barðdal & Gildea, 2015) would go along with an increased emergence of connections and their subsequent strengthening.

To summarize the main point of this section, I consider it an open question whether the distinction of constructionalization and constructional change should perhaps be complemented by another way of differentiating between types of change in the constructional network, and I suggested that dividing up types of change along the parts of the network they affect might be a useful alternative. Notably the differentiation between node changes and connectivity changes seems to hold some promise. All of these types would still be subsumed under the heading of constructional change, which signifies that the changes affect individual form-meaning pairs and their connections (cf. Hilpert, 2013, p. 13), and not the language system as a whole.

4. Are there nodes in the constructional network?

The preceding section presented a view of the constructional network in which constructions form nodes that are interlinked through connections, and it also proposed a distinction of types of change that pertain either to the nodes or to the connections between them. As a result, the heading of this section may seem like a rhetorical question, but importantly, it is not. Whereas Construction Grammar very much embraces the idea that knowledge of language is a network, and nothing else in addition (Hilpert, 2014, p. 22), it has to be conceded that most information that is supposed to be stored in that network is stored in the nodes. A construction such as *What's X doing Y* is envisioned as a bundle of features with values (Kay & Fillmore, 1999, p. 22) that speakers mentally represent. Through inheritance links and subpart links (cf. Hilpert, 2014, Chapter 3), each construction is connected with other parts of the network. If constructions are seen as nodes in a network, it follows that these nodes can have substantial internal structure and complexity. This contrasts with many other approaches to networks, which tend to limit the internal complexity of nodes to a threshold of activation

and, when that threshold is exceeded, a subsequent response in which the node 'fires', thereby sending activation to other nodes (Gurney, 1997). Constructions are much more complex than that. So whereas in Construction Grammar, most of linguistic structure is actually represented in the nodes, other approaches try to model linguistic knowledge by representing it exclusively as a configuration of connections. This difference is pinpointed by Hudson (2015, p. 692):

> I believe that language is, indeed, a network, and that this network is, indeed, a structure. Many other readers may protest that they too see language as a network; after all, cognitive linguists envisage 'an elaborate network comprising any number of conventional units linked by categorizing relationships' (Langacker 2000: 12) or a 'network of constructions [which] captures our grammatical knowledge of language in toto, i.e. it's constructions all the way down' (Goldberg 2006: 18). But notice that in these cases the complex units which the network connects have their own internal structure which is not part of the network. Similarly, there are network-based theories of individual areas, such as Network Morphology (Brown and others 1996, Corbett and Fraser 1993, Fraser and Corbett 1997), and there are network-based theories of how usage affects storage (Bybee and Beckner 2010). But network theory goes further by claiming that 'it's networks all the way down'.

Whether Langacker's position is correctly represented in this quote is a matter of debate, since Langacker (2005, p. 107) has stated quite clearly that in his Cognitive Grammar, the only structures that are posited are sounds and meanings, while links (symbolic associative links, categorizing inheritance links, syntagmatic associative links) are responsible for all the remaining work. Yet, Hudson's characterization adequately describes a broad consensus in those branches of Construction Grammar that are not quite as reductionist as Langacker's Cognitive Grammar.

What then is undesirable about nodes with internal complexity? A network architecture with simple nodes, in which the structure lies mainly in the connections, has a very strong appeal for two reasons. First, such an architecture could be mapped more or less directly onto neurophysiological processes, thereby creating a bridge between the study of language and the study of the human brain. A second reason is the relative ease with which such networks can be computationally implemented. If a network lends itself to computational simulations, its empirical predictions can be put to the test by manipulating its parts and studying the resultant differences in its behavior (cf. Wellens et al., 2013; van Trijp, 2010, 2012). Schmid (2017, p. 25) actually envisions a Construction Grammar in which, lining up with Hudson's own views, associative connections are the only mental representations of linguistic knowledge:

> Usage-based models usually assume that entrenchment operates over constructions and constructional schemas which are characterized as form-meaning pairings. Furthermore, they claim that these constructions and schemas are related to each other in a massive associative memory network organized mainly in terms of hierarchical relations […]. The present proposal diverges from this idea in two important ways: First, it rejects the distinction between constructions serving as nodes in the network and relations between nodes and instead assumes that linguistic knowledge is available in one format only, namely associations.

It is not the aim of this paper to argue that Diachronic Construction Grammar should completely give up the current implicit consensus and with it any discussion of constructions as nodes in a network. At the same time, there are clear benefits to a model of linguistic knowledge in which greater emphasis is placed on connections, especially when change in the network is of central concern. As Hilpert & Diessel (2017, p. 70) argue, a connection-centered view of linguistic knowledge and change in that knowledge allows for an inherently dynamic representation of the constructional network. When connections spread activation and, as a result, change in strength, those are dynamic processes. Theoretical models exist that offer explicit descriptions of these processes (Gurney, 1997). We can thus add to the advantages mentioned above that a focus on connections brings us closer to a theory that models knowledge of language not as a static inventory of regularities, but rather as a fluid system which incorporates change as an inherent characteristic. This is a perspective that is also voiced by von Mengden & Coussé (2014, p. 14), who call for a view of "modifiability and adjustability as part of the nature of linguistic structures and hence, of the linguistic system itself".

Importantly, the two perspectives are not mutually exclusive. With regard to actual analyses of historical language data, the same observation can often be couched in terms of either one or the other. To take an example, Colleman (2015) presents an analysis of semantic and collocational change in the Dutch *krijgen*-passive. The empirical observation is that as the 20th century unfolds, the construction is used with an increasingly wide range of verbs in its participle slot. If the node view is given priority, the development of the construction could be described as an extension of its meaning pole or as a deterioration of its semantic constraints. Alternatively, the same phenomenon could be viewed as the formation of associative connections. New symbolic links emerge between the form of the *krijgen*-passive and new meanings, to which the form was not connected during earlier periods, and new syntagmatic links emerge between *krijgen* and participle types that earlier generations of speakers did not usually combine with it. The two descriptions capture the same insight, but highlight different aspects.

It is also worth noting that some work on a more connection-centered Construction Grammar is already underway. An example of this is Torrent (2015),

who develops an analysis of infinitive constructions in Brazilian Portuguese in order to propose two hypotheses about how constructional networks may change. The first of these is labeled the Constructional Convergence Hypothesis. It claims that new links can emerge between constructions that are historically unrelated, but that share aspects of form and come to serve similar functions. The second hypothesis that Torrent advances is called the Construction Network Reconfiguration Hypothesis, which is concerned with the changes in connectivity that result from the emergence of new constructions. As new constructions arise, the network readjusts itself and new inheritance links are created. In a similar vein, Pijpops & Van de Velde (2016) study an effect that reveals how connections between constructions influence speakers' language use. This effect, which they call constructional contamination, can be observed when a frequent structural pattern in one construction influences the usage of another, thereby 'contaminating' that construction. Pijpops & Van de Velde illustrate this concept with variation in the Dutch partitive genitive construction, which is exemplified by phrases such as *iets verkeerds* 'something wrong', in which the adjective is marked with an *-s* suffix, or *iets leuk* 'something fun', in which the bare form of the adjective is used. Pijpops & Van de Velde show that a preference for the *s*-less variant in the partitive genitive construction can be predicted by the frequency of co-occurrence of the quantifier and the adjective in other constructions that, as a matter of default, do not have the s-suffix. If uses of for example *iets verkeerd* in contexts such as *dat iets verkeerd geïnterpreteerd wordt* 'that something gets wrongly interpreted' are frequent enough, they will lead speakers to adopt the s-less pattern even in the partitive genitive construction, where it is an option. While both Torrent (2015) and Pijpops & Van de Velde (2016) focus on changing connections between pairs of constructions, it is also useful to consider connectivity change between a single grammatical construction and its many lexical collocates. Hilpert (2016) presents a diachronic analysis of the shifting collocational preferences of the English modal auxiliary *may*, which changes over time with regard to the lexical verbs that typically co-occur with it. As time goes on, certain lexical verbs become more and more attracted to *may*, and these shifting patterns of connectivity reflect the secondary grammaticalization of *may* from a deontic modal auxiliary into a modal that is primarily associated with epistemic meaning.

In summary, the study of connectivity changes in Diachronic Construction Grammar has already gained some momentum, and more interesting results can be expected to follow in the future, as this area of study is further developed.

5. Concluding remarks

Diachronic Construction Grammar is a research program that is currently under development, and that as a consequence still has to form a common core of assumptions and a shared vocabulary to talk about the phenomena that are of interest to its community of researchers. It was the aim of this paper to raise questions about three aspects of Diachronic Construction Grammar that merit some reflection. Specifically, this paper discussed the status of constructions as mental representations of language, the concept of constructionalization, and the relative importance of nodes and connections in the constructional network. With regard to the first issue, it can be stated that in the community there is currently no shared commitment to the goal of investigating mental representations, rather than linguistic forms that can be described in relative independence of cognitive matters. It was argued here that there is room for all kinds of constructionist analyses, be they concerned with cognition, linguistic structure, or other aspects of language such as interaction or social practice. It would, however, be useful for researchers to indicate explicitly how their work is situated in the multi-faceted enterprise of Diachronic Construction Grammar.

The second question addressed existing typologies of processes of change, and in particular the notion of constructionalization (Traugott & Trousdale, 2013), which captures the emergence of a new node in the constructional network. The example of syntactic and semantic change in the English verb *confirm* was presented in order to argue that the distinction between constructional change and constructionalization is in some cases problematic, and that constructionalization can only ever be identified in relation to a starting point that is chosen by the analyst. It was suggested that an alternative way to distinguish different types of change in the constructional network could be to ask which parts of the network are affected by a given process of change. In this context, changes in connections were identified as a particularly relevant phenomenon that has, up to now, not received all the attention it deserves.

The third issue that was brought up in this paper concerned the overall architecture of the constructional network, specifically the place that nodes have in that network vis-à-vis the connections that link them with each other. Diachronic Construction Grammar shares with its synchronic counterpart a consensus view in which the lion's share of speakers' linguistic knowledge is represented within the nodes, which thus have complex internal structure. The discussion in this paper called for a shift in perspective, arguing that a connection-centered view of linguistic knowledge would allow for closer ties with psychological and neuro-linguistic work, better options for computational implementation, and most importantly a more realistic view of linguistic knowledge as an inherently dynamic

system. There is a growing awareness for these issues, which is exemplified by publications such as Hundt et al. (2017). It will be an important goal to understand how processes of language change are ultimately related to language processing in the here and now.

While these three questions can be discussed independently of each other, it is clear that they are in fact intimately related. Questions about constructions as mental representations directly lead to questions about the network structure that is assumed to contain these constructions, which further lead to questions about how the network can change over time and what types of change we expect to see in it. How one of these questions is answered has consequences for how the other questions can be approached. Working out a coherent set of answers is beyond the scope of any single paper, but luckily there is a growing, vibrant community of researchers in Diachronic Construction Grammar, so that this task can be tackled collectively. Hopefully, the points raised in this paper will contribute in some way to this process.

References

Barðdal, J., & Gildea, S. (2015). Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In J. Barðdal, E. Smirnova, S. Gildea, & L. Sommerer (Eds.), *Diachronic Construction Grammar* (pp. 1–150). Amsterdam: John Benjamins.

Barðdal, J., Smirnova, E., Gildea, S., & Sommerer, L. (Eds.). (2015). *Diachronic Construction Grammar*. Amsterdam: John Benjamins.

Bergs, A., & Diewald, G. (Eds.). (2008). *Constructions and Language Change*. Berlin: de Gruyter.

Börjars, K., Vincent, N., & Walkden, G. (2015). On constructing a theory of grammatical change. *Transactions of the Philological Society*, 113(3), 363–382.

Brown, D., Corbett, G. G., Fraser, N. M., Hippisley, A., & Timberlake, A. (1996). Russian noun stress and network morphology. *Linguistics*, 34, 53–107.

Bybee, J. L. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The New Psychology of Language*, Volume II (pp. 145–167). New Jersey: Lawrence Erlbaum.

Bybee, J. L. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, J. L., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Bybee, J. L., & Beckner, C. (2010). Usage-based theory. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 827–855). Oxford: Oxford University Press.

Colleman, T. (2015). Constructionalization and post-constructionalization: The constructional semantics of the Dutch *krijgen*-passive in a diachronic perspective. In J. Barðdal, E. Smirnova, S. Gildea, & L. Sommerer (Eds.), *Diachronic Construction Grammar* (pp. 216–258). Amsterdam: John Benjamins.

Corbett, G. G., & Fraser, N. M. (1993). Network morphology: A DATR account of Russian nominal inflection. *Journal of Linguistics*, 29, 113–142.
Davies, M. (2010). *The Corpus of Historical American English: 400 million words, 1810–2009*. http://corpus.byu.edu/coha/.
De Smet, H. (2013). *Spreading patterns: Diffusional change in the English system of complementation*. Oxford: Oxford University Press.
Evans, V. (2016). Cognitive Linguistics. In S. E.F. Chipman (Ed.), *The Oxford Handbook of Cognitive Science* (pp. 283–299). Oxford: Oxford University Press.
Fillmore, C., Kay, P., & O'Connor, C. (1988). Regularity and Idiomaticity in Grammatical Constructions: The Case of *let alone* . *Language*, 64, 501–538.
Fraser, N. M., & Corbett, G. G. (1997). Defaults in Arapesh. *Lingua*, 103(1), 25–57.
Goldberg, A. E. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219–224.
Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
Gurney, K. (1997). *An Introduction to Neural Networks*. London: Routledge.
Heine, B. (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
Heine, B., Narrog, H., & Long, H. (2016). Constructional change vs. grammaticalization: From compounding to derivation. *Studies in Language*, 40(1), 137–175.
Hilpert, M. (2011). Dynamic visualizations of language change: Motion charts on the basis of bivariate and multivariate data from diachronic corpora. *International Journal of Corpus Linguistics*, 16(4), 435–461.
Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word-formation and syntax*. Cambridge: Cambridge University Press.
Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.
Hilpert, M. (2015a). Historical Linguistics. In E. Dabrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 346–365). Berlin: De Gruyter Mouton.
Hilpert, M. (2015b). From hand-carved to computer-based: Noun-participle compounding and the upward-strengthening hypothesis. *Cognitive Linguistics*, 26(1), 1–36.
Hilpert, M. (2016). Change in modal meanings: Another look at the shifting collocates of *may* . *Constructions and Frames*, 8(1), 66–85.
Hilpert, M., & Diessel, H. (2017). Entrenchment in construction grammar. In H.-J. Schmid (Ed.), *Entrenchment, Memory and Automaticity. The psychology of linguistic knowledge and language learning* (pp. 57–74). Berlin: Mouton de Gruyter.
Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Second edition. Cambridge: Cambridge University Press.
Hudson, R. (2015). Review of Rolf Kreyer, The nature of rules, regularities and units in language: A network model of the language system and of language use. Mouton De Gruyter, 2014. *Journal of Linguistics*, 51(3), 692–696.
Hundt, M., Mollin, S., & Pfenninger, S. E. (Eds.). (2017). *The Changing English Language: Psycholinguistic Perspectives*. Cambridge: Cambridge University Press.
Janda, R. D. (2001). Beyond 'pathways' and 'unidirectionality': on the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences*, 23(2), 265–340.

Kay, P., & Fillmore, C. (1999). Grammatical Constructions and Linguistic Generalizations: The *What's X Doing Y?* Construction. *Language*, 75(1), 1–33.

Langacker, R. (2000). A dynamic usage-based model. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). Stanford, CA: CSLI Publications.

Langacker, R. (2005). Construction Grammars: Cognitive, Radical, and Less So. In F. J. Ruiz de Mendoza Ibáñez, & M. S. Peña Cervel (Eds.), *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction* (pp. 101–159). Berlin: Mouton de Gruyter.

Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language*, 14(2), 177–202.

Nørgård-Sørensen, J., & Heltoft, L. (2015). Grammaticalisation as paradigmatisation. In A. D. M. Smith, G. Trousdale, & R. Waltereit (Eds.), *New Directions in Grammaticalization Research* (pp. 261–292). Amsterdam: John Benjamins.

Perek, F. (2015). *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Amsterdam: John Benjamins.

Petré, P. (2014). *Constructions and environments: Copular, Passive and related Constructions in Old and Middle English*. Oxford: Oxford University Press.

Pijpops, D. & Van de Velde, F. (2016). Constructional contamination. What is it and how do we measure it? *Folia Linguistica*, 50(2), 543–581.

Schmid, H.-J. (Ed.). (2017). *Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge*. Boston: APA/Walter de Gruyter.

Sommerer, L. (2015). The influence of constructions in grammaticalization: revisiting category emergence and the development of the definite article in English. In J. Barðdal, E. Smirnova, S. Gildea, & L. Sommerer (Eds.), *Diachronic Construction Grammar* (pp. 109–140). Amsterdam: John Benjamins.

Torrent, T. (2015). On the relation between inheritance and change: The Constructional Convergence and the Construction Network Reconfiguration Hypotheses. In J. Barðdal, E. Smirnova, S. Gildea, & L. Sommerer (Eds.), *Diachronic Construction Grammar* (pp. 175–214). Amsterdam: John Benjamins.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Traugott, E. C. (2015). Toward a coherent account of grammatical constructionalization. In J. Barðdal, E. Smirnova, S. Gildea, & L. Sommerer (Eds.), *Diachronic Construction Grammar* (pp. 51–79). Amsterdam: John Benjamins.

van der Auwera, J., Van Olmen, D., & Dumon, D. (2015). Grammaticalization. In E. Dabrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 634–650). Berlin: De Gruyter Mouton.

von Mengden, F., & Coussé, E. (2014). The role of change in usage-based conceptions of language. In E. Coussé, & F. von Mengden (Eds.), *Usage-based approaches to language change* (pp. 1–19). Amsterdam: John Benjamins.

Van Rompaey, T., Davidse, K., & Petré, P. (2015). Lexicalization and grammaticalization: the case of the verbo-nominal expressions *be on the/one's way/road* . *Functions of Language*, 22(2), 232–263.

van Trijp, R. (2010). Grammaticalization and Semantic Maps: Evidence from Artificial Language Evolution. *Linguistic Discovery*, 8(1), 310–326.

van Trijp, R. (2012). The Evolution of Case Systems for Marking Event Structure. In L. Steels (Ed.), *Experiments in Cultural Language Evolution* (pp. 169–205). Amsterdam: John Benjamins.

Wellens, P., van Trijp, R., & Beuls, K. (2013). Fluid Construction Grammar for Historical and Evolutionary Linguistics. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 127–132.

Wolk, C., Bresnan, J., Rosenbach, A., & Szmrecsanyi, B. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica*, 30(3), 382–419.

Verb constructions

Grammaticalisation cut short

A diachronic constructional view on English posture verbs

Mégane Lesuisse & Maarten Lemmens
Université de Lille | UMR STL 8163, France

This article provides a corpus-based diachronic constructional analysis of the three English cardinal posture verbs (hence CPVs) *sit*, *stand* and *lie*. Contrary to their equivalents in genetically related languages, the English CPVs have not grammaticalised into locative, progressive, or habitual markers or into copulas. The data reveal some degree of grammaticalisation in the earlier stages which suggests that the grammaticalisation was cut short. The paper evaluates Newman's (2009) hypothesis that the English CPVs did not further grammaticalise because of their increased use to refer to a dynamic event of getting into a sitting, standing or lying posture. The data also reveal that two (more strongly grammaticalised) copula constructions lingered on much longer, thereby constituting grammaticalised 'islands', especially in idiomatic uses.

Keywords: posture verbs, semantic bleaching, grammaticalisation, copular constructions, construction grammar

1. Introduction

This article analyses the diachronic development of English cardinal posture verbs (henceforth, CPVs) *sit*, *stand* and *lie*.[1] These three verbs are said to be cardinal since they refer to the three basic postures of human beings. In many unrelated languages, these verbs show considerable semantic extension and grammatical-

1. This paper was first presented at the Workshop "Grammaticalisation meets Construction Grammar" organized at the University of Gothenburg, Sweden, by Peter Andersson, Evie Coussé and Joel Olofsson, whom we also thank for their constructive feedback on the earlier version of this paper. We also thank the reviewers of this paper for their insightful comments. Financial support of the Institute of Modern Languages (ILT) at the University of Leuven, Belgium for the workshop participation is hereby also gratefully acknowledged.

https://doi.org/10.1075/cal.21.c3

isation patterns (cf., among others, Kvist Darnell, 2008; Serra Borneto, 1996; Lemmens, 2002, 2005a; Newman, 2002; Newman & Rice, 2006; Van Oosten, 1984). Newman & Rice (2006) provide a list of all the outputs of grammaticalisation of 'sit', 'stand' and 'lie' attested across languages:

SIT: progressive marker, present marker, habitual marker, copula, locative marker, concessive, counter-expectation marker ('not yet').
STAND: (stative) progressive marker, durative marker, present marker, ingressive marker, copula, distributive marker, (motion) consecutive clause marker
LIE: habitual marker, progressive marker.

As Newman & Rice point out, what is fascinating about CPVs is that "they enter into constructions which seem to have very little to do with either static configuration of a theme or its dynamic positioning" (2006, p. 352) and, in addition to referring to location, they have come to express more functional relations such as existence, tense/aspect, deixis, voice, classification, or social status (Newman, 2002).

The Germanic languages illustrate such grammaticalised uses quite nicely. For instance, in Dutch and the Scandinavian languages (but a bit less in German), the CPVs are used as basic locative markers for expressing the location of concrete entities in space, e.g., CUP (STAND) ON TABLE; CLOTHES (LIE) ON GROUND (cf., e.g., Lemmens, 2002 on Dutch; Kvist Darnell, 2008 on Swedish; or Serra Borneto, 1996 on German). Kuteva (1999) claims that the use as basic locative verbs is a precondition (or a first step) for further grammaticalisation (to markers of progressive, durative or habitual aspect) to take place. This view is quite plausible, since in the Germanic languages, the CPVs are indeed used as (semi-) auxiliaries to express the progressive aspect of an event, particularly in Dutch, e.g., *Ik zit maar wat rond te lopen* 'I just sit to walk around' = 'I am walking around'; (cf. Lemmens, 2005b; 2015; 2017). However, English seems to be the odd one out in the family of Germanic languages, as the English locative construction prefers the general verb *be* (<X BE $PREP_{locative}$ Y>, e.g., *The cup* ***is*** *on the table*) and also the progressive uses *be*.[2] The special status of English in this regard is also observed by Dunn et al. (2007). While their article is primarily concerned with locative verbs in four languages unrelated to the Germanic family (Saliba, Tiriyo, Chukchi and Lavukaleve), at the end of their article they mention the difference between Dutch and English:

> In fact, if one compares English *sit*, *stand*, *lie*, and *hang* to […] Dutch *zitten, staan, liggen*, and *hangen*, one is struck by the fact that the two sets of verbs are cognate

2. See also Lemmens (2005b) on inter-Germanic differences in the use of posture verbs, relating them to motion verbs.

> and very close in meaning, yet the Dutch set forms a real positional verb system, while the English set does not. It may be that languages like English, Saliba and Lavukaleve have a potential positional verb system, while Dutch [has] an actual positional verb system. (Dunn et al., 2007, p. 189)

Similarly, Lemmens (2005a) has suggested a continuum of posture verbs usage with Dutch on the one end (high obligation to use a CPV for a locative event) and English on the other (low obligation), with German and Swedish straddling the middle.

The question left unanswered (by both Dunn et al., 2007; and Lemmens, 2005a) is why there is this difference between English and the other Germanic languages as far as the use of CPVs is concerned (a potential vs. an actual positional system) and where it comes from. This is the main research question that we address in this article. To answer that question, we investigated the use of CPVs in the entire Modern English period (1500–1920), considering three types of constructions in which CPVs occur:

1. The ***postural construction*** , i.e. *Subject + CPV + (particle) + (locative complement)*, such as instantiated by *We sat down (at the feast)*. As the name suggests, this usage refers to the posture of a human Figure where the expression of the Ground (i.e., a locational complement) is optional.
2. The ***locational construction*** , i.e., *Subject + CPV + locative complement*, e.g., *The student's desk stood by the window.* The Figure is the entity (animate or inanimate) being located in respect to a second entity, the Ground which, like the Figure, could be either concrete or abstract. If either the Figure or the Ground is abstract, then we talk about "metaphorical location", as illustrated by *Dismay sat in every countenance*.
3. The ***copular construction*** , which can either be nominal (*Subject + CPV + Noun*, e.g., *[He] stood Godfather to my little Girl*) or adjectival (*Subject + CPV + Adjective*, e.g., *Convention Committees sit supreme over them*).

The construction helps us to decide which use of the CPV it is (postural, locational, metaphorical, copula-like use), even if there is no systematic correspondence and a particular construction cannot be exclusively associated with one particular use of the CPV use. However, as we will see, such a constructional view gives a more accurate and a more nuanced view on the (non-)grammaticalisation of the CPVs compared to an analysis which exclusively hinges on the semantics of the CPV itself. The three constructions listed above can actually be seen as building a cline of grammaticalised use, the postural construction being a (non-grammaticalised) use of the CPV, the locational uses a first step in the grammaticalisation process (semantic bleaching). As Dunn et al. (2007) point out,

the development of CPVs into locative markers exclusively depends on the prototypical anthropocentric uses. The copular constructions can be considered as the most grammaticalised, since the lexical verb has been reanalysed as a copula. Logically, such a copular reading (e.g., *The house stood empty* vs. *The house was empty*) is only possible if the posture verb has established as a basic locative verb (in our example, as the basic verb for expressing the location of a house). Adopting a constructional view allows us to measure what Himmelmann (2004, p. 32) has called "host-class expansion" of English CPVs; paraphrasing his definition, we thus consider the class of elements the CPV is in construction with. Coussé (forthcoming) calls this a "category-internal change". The framework of Construction Grammar in which our analysis is situated contributes to a better understanding of such category internal changes within and across these constructions. Our analysis comprises both a quantitative and qualitative perspective. The former consists in analysing the productivity of a given construction, via methods that have been established in the Construction Grammar framework for other constructions (see Section 4 below); this degree of productivity can then be taken as an indication of the degree of host-class expansion and thus of its degree of grammaticalisation. This would allow us, for example, to evaluate Kuteva's (1999) claim mentioned above that the grammaticalisation of CPV (e.g. to aspectual markers) is preceded by their becoming established as basic locative verbs used as the obligatory expression for the location of inanimate entities. The qualitative analysis, in turn, leads to a nuanced picture overall, showing that some uses displayed high degree of grammaticalisation which also lingered on in the language for a longer time as grammaticalised 'islands'.

Our corpus-based study, investigating the use of the CPVs in the entire Modern English period (1500–1920), considers two more specific questions:

(1) What was the degree of grammaticalisation of the English CPVs, especially in the earlier period? Based on OED data, Lemmens (2014) has shown that the English CPVs *were* used in patterns largely similar to those still found in Modern Dutch for example, but further confirmation of their entrenchment (at that time) is still needed, which is what we aspire to offer here.
(2) Can we find evidence for the demise of the grammaticalised uses of the English CPVs, notably when and why it occurred? In particular, we will compare the frequency of locative and copular uses of CPVs to that of the dynamic (postural) uses of these verbs, which Newman (2009) hypothesizes to be responsible for the demise.

In line with Himmelmann (2004), our study considers the semantic bleaching and (possible) grammaticalisation of CPVs from a more constructional perspective, where we not only take the syntagmatic environment into account, but also

consider what larger effect it has on the grammaticalisation process as a whole. More particularly, via standard corpus linguistic methods, we evaluate the impact of the particle-less construction with CPVs (referring to a dynamic situation) as the trigger of the demise of these verbs and the cutting short of their further grammaticalisation.

The article is organized as follows. First, in Section 2, we will sketch the uses of the English CPVs in earlier periods, describing their evolution from postural to locational verbs, a form of semantic extension or bleaching (Section 2.1), to even copular uses (Section 2.2); we also discuss the three hypotheses that have been suggested by Lemmens (2014) to account for their demise (Section 2.3). Next, in Section 3, we will describe the corpora on which our study is based and the method of analysis. In Section 4, we present our main findings with regard to their degree of grammaticalisation, notably by investigating two copula-like constructions (CPV+Adjective, CPV+Noun) and their further diachronic development. As we will show, this is the story of a grammaticalisation cut short by language internal changes.

2. The unexpected path of English Posture Verbs

2.1 *From posture to location and beyond*

Unlike the other Germanic languages, Present-Day English does not use posture verbs as basic location verbs, but generally prefers the neutral verb *be* (e.g., *My car **is** in the garage* rather than *My car {**stands/is standing**} in the garage*). On the basis of data from the *Oxford English Dictionary*, Lemmens (2014) however shows that in earlier periods, the English CPVs were used in contexts that are quite similar to those one still finds in contemporary Dutch or Swedish.[3] More specifically, from a qualitative perspective the English CPVs were amply used in locative contexts (i.e., uses that refer to the location of an inanimate entity) as is still common in contemporary Dutch, for example. We cannot give a full overview of all these older uses here, but restrict ourselves to the major locative usages (and their motivations).

For example, in earlier stages of English, *stand* is used to refer to objects placed on their base (analogous to human feet) and thus also in a functional

3. Lemmens (2014) focuses primarily on the uses of *stand*; the present discussion also considers *lie* and *sit*. The present discussion is kept quite short and no quoted examples are given (either from the OED or the corpus); the reader is referred to Lemmens (2014) for some examples illustrating these and other uses.

position. This use of *stand* (as opposed to *lie*) can often be explained in terms of BE ON ONES BASE, where the base is the origin of a mental vertical scanning operation (cf. also Serra Borneto, 1996). The idea of an entity resting on its base also motivates the (older) use of *stand* to refer to plants that grow upward from their roots (=base), e.g., trees or crops 'standing' in the field.[4] The mental scanning 'away from the base' that underlies such uses, allows an image schematic transformation, i.e., rotation, where the mental scanning can be in any direction (i.e., not just vertical), such as branches 'standing' on the trees, or a moustache 'standing' sideways.

The semantic extension from 'human being on feet' to 'object on base' to 'scanning away from origin' can be represented as in Figure 1 (reproduced from Lemmens, 2014):

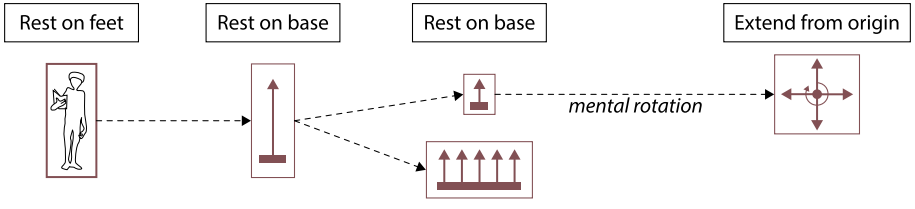


Figure 1. Simplified schematic network for *stand*

Another extended use of *stand* is its reference to a human or any other entity being in its functional position, a usage that can be related back to a human on its feet and thus being in its most functional position (as can also be said to be the case for objects on their base). A fleet could thus be said be standing at sea, referring to its readiness to sail out or attack.[5]

In short, central to a coding with *stand* is the presence of a base and the functional position that is associated with it. The real dimensions come into play only in the absence of a base, for entities that display a (salient) difference in height and width. A book, for example, does not really have a base; if it is in an upright position (on its smallest side, like in a bookshelf) *stand* could (can) be used, whereas *lie* will be used if it is on its front or its back. Strikingly, in Present-Day English the verb *sit* rather than *stand* is often used to express the location of objects on their base (Newman, 2002 calls this "inactivity *sit*") where Dutch would use *staan*. English *stand* could be used in this context; however, Newman considers such uses of *stand* to be "rather stilted, more literary than colloquial, a little pretentious almost" (Newman, 2002, p. 9). As said, overall Present-Day English does not use the posture verbs any more but prefers the neutral verb *be*.

4. Most likely, this usage underlies Present-Day English *a stand (of trees)*.

5. See also Gibbs et al. (1994) for such more metaphorical uses of (Present-Day) English *stand*.

The contemporary use of *sit* illustrated above is, in fact, quite different from its earlier usages. For example, the verb used to be used refer to the location of smaller animals such as frogs, birds, insects and serpents. Strikingly (again), these uses find a perfect equivalent in contemporary Dutch where *zitten* 'sit' is to be used. The posture of birds and frogs (and rabbits and mice) arguably resembles that of a crouched human being (=sitting with legs bent); insects and legless creatures are close to the ground and can therefore be coded with *sit* (if they are said to be lying, they are dead).

One of the common extended uses of *sit* one finds in the older English data (but also, significantly again, in contemporary Dutch) is where it refers to 'being in a fixed position' or 'being (closely) contained', like a ring said to sit on one's finger, a post sitting snugly in their place, or even metaphorically, a spirit 'sitting' in one's eyes. These uses are motivated by a (metonymic) extension (i) from the (partial) containment when sitting (snugly) in a chair to the larger space or (ii) from the close contact with the chair's surface to mere contact, as represented in Figure 2.

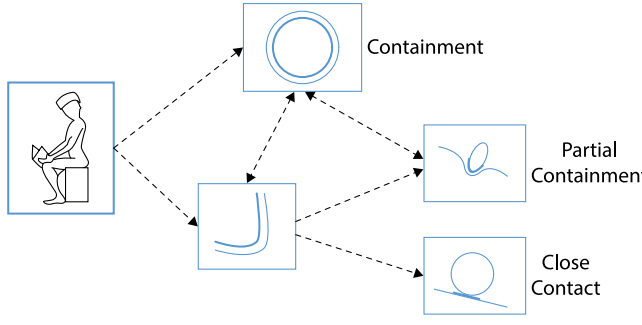


Figure 2. Metonymic extension for CONTAINMENT *sit*

Finally, also the verb *lie* could be used in a wider range of contexts than in contemporary English (even if here more uses seem to have been preserved, cf. also below). In addition to the use for entities not on their base (typically but not necessarily having an orientation that is more horizontal than vertical) already mentioned above, the verb was used (as diagrammed in Figure 3) to express the location (i) of round or symmetrical objects; (ii) of saliently elongated or horizontally extending entities like roads, bridges, and metaphorical extensions thereof, such as lines or frontiers; and (iii) of substances that readily assume a horizontal extension under the force of gravity.

That round or symmetrical objects are coded with *lie* finds its explanation in the absence of dimensional differentiation (cf. Serra Borneto, 1996; hence the label 'DIM-LESS' in Figure 3); in other words, there is no opposition between vertical and horizontal (except when they have a base and functional position, which

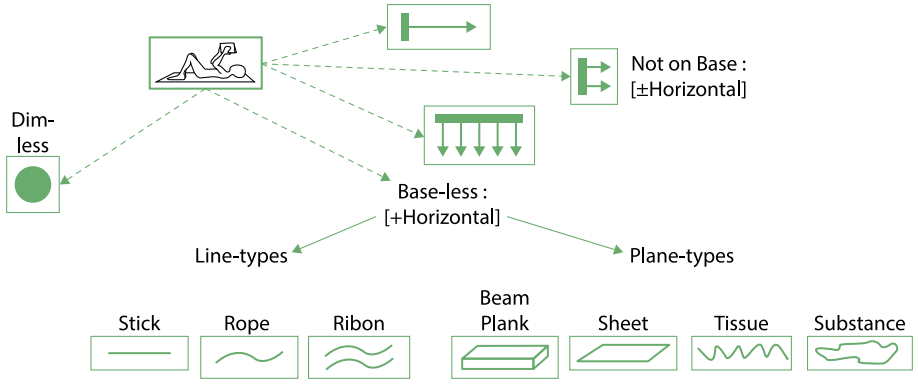


Figure 3. Semantic extensions for *lie*

triggers a coding with *stand* if resting on their base). The absence of dimensional differentiation may also be the explanation for why *lie* is used to talk about abstract entities (a usage that continues till the present-day, cf. below), such as choices, causes or explanations that are said to 'lie'.

Another use of *lie* that was common in earlier periods is in reference to the locations of cities, villages and the like where the verb designates their horizontal extension, basically a two-dimensional ('sheet-like') view of the location; Serra Borneto (1996) has termed this usage "geotopographical location" (cf. also Lemmens, 2002; Lemmens & Perrez, 2010 for examples of this use for Dutch *liggen* 'lie').

The short description of the semantic extensions of the CPVs in older periods of English shows that they had locative uses that not only went much further than in present-day English but which also have striking similarities with what one still finds in the other Germanic languages (especially in Dutch). As the next section will show, they also developed copular uses.

2.2 *From location verb to existential copula*

As already mentioned by Newman & Rice (2006), in several unrelated languages, the verb STAND has grammaticalised into a copula and/or existential marker. In the Germanic languages, for instance, existence is expressed through prefixed forms of STAND, e.g., Dutch *bestaan* (from *staan*), German *bestehen* (from *stehen*), and Swedish *bestå* (from *stå*). Our historical corpora reveal traces of such a grammaticalisation for English *stand*, as in *[He] stood Godfather to my little Girl that is now in her Mother's Arms. (1710–1780).* This situation could also have been expressed by the neutral verb *be* (preferred coding in present-day English). Under the assumption that any difference of form implies a difference of conceptualiza-

tion, Lemmens (2014) states that the use of *stand* as an existential copula emphasizes prolonged existence, whereas *be* expresses mere location and existence.

From a constructional point of view, the copular construction can be summarized as a form-meaning pairing, which Petré (2012) characterizes for English as a [*SubjNP*$_{NOM}$ + *Copular Verb* + *SubjCompXP*$_{NOM}$] constructional schema; the subject complement can be introduced by a preposition (but this is not always so) and is formed of an adjective or a noun, both in the nominative case. Petré defines the meaning of the Copular Construction as follows: "The construction predicates a non-eventive Subject Complement of a non-agentive Subject, and adds tense and aspect to it through Copular Verb" (Petré, 2012, p. 30). The qualifying adjectives occurring as the Subject Complement conform to the idea that the Copular Construction does not designate an action but a state or change of state. Concerning the Copular verb itself, Petré argues that it must have only one participant in terms of the event structure it evokes when it is *not* used in the Copular Construction. In other words, for a verb to fit well in the copular verb slot, it has to have the potential to be used intransitively. Consequently, verbs of (intransitive) motion, growth or position are perfect candidates. While the posture verb *stand* is much more common in this use in (older) English, copula-like patterns have also been attested in our Modern English corpora (see Section 3 below) for *sit* and *lie*.

Petré distinguishes three subtypes of the more schematic Copular construction given above: (i) the ***Nominal Copular Construction*** [*SubjectNP + CPV + SubjCompNP*], as in Example (1a); (ii) the ***Adjectival Copular Construction*** [*SubjectNP + CPV + SubjCompAdjP*], illustrated in (1b); and (iii) the ***Prepositional Complement Construction*** [6] where the Subject Complement is introduced by a preposition, as in (1c).

(1) a. I *stood* a calm spectator of the flames. (*stand*, 1710–1780)
b. You *stand* amazed at this preamble. (*stand*, 1710–1780)
c. I *sit* as a criminal judge on acts of his. (*sit*, 1710–1780)

As indicated earlier, we consider these Copular Constructions as the most grammaticalised patterns attested in English as they build on the basic locative uses having been established; as we will show below, they do provide an interesting test case to trace the evolution of English CPVs throughout the Modern English period (1500–1920).

6. This last construction is to be contrasted with the locational construction analysed in the previous section which presents two participants, the Figure –participant 1- being located in space in respect to the Ground –participant 2-. Petré states that these similarities are due to the fact that prepositional copular constructions probably develop out of locational constructions. (Petré, 2012, p. 32)

While the presentation above does not provide a full account of the rich semantic extensions of the English CPVs in earlier periods, it clearly emerges that they were used in ways that sound quite unfamiliar (or even plainly unidiomatic) to a speaker of contemporary English, while they sound perfectly natural to a speaker of contemporary Dutch. In other words, these data clearly suggest that in these earlier periods English was most likely not all that different from the other Germanic languages in its use of CPVs as basic locative verbs used to express the location of any entity in space as well as in metaphorical extensions thereof, and even in copular uses.

Even if the OED data on which Lemmens (2014) is based quite convincingly reveal the more extended and even grammaticalised use of the English CPVs, further *quantitative* support is needed to confirm their entrenchment as basic locative verbs; this is precisely what we offer in Section 4 below. However, before turning to these results, we present the hypotheses that have been suggested to explain the demise of the English CPVs.

2.3 Possible causes for the demise of English posture verbs

If it is indeed the case that the English CPVs had more extensive uses in earlier periods, where especially the locative uses stand out to the contemporary speaker, then English must somehow have lost these along the way. Lemmens (2014) has suggested three different explanations for the 'demise' of the English CPVs.

The first hypothesis is that the posture verbs have disappeared under French influence. As is commonly known, English has borrowed massively from French, including verbs relating to either motion or location, such as *enter*, *exit*, *place*, *situate*, etc. These verbs do not fit the Germanic pattern very well, since they express the path (or location) in the verb when Germanic language typically express the path in a satellite (hence, Talmy's 2000 characterisation of English as a satellite-framed language) and the manner of motion or location is typically expressed in the verb. This first hypothesis is quite unlikely, as it would only have affected the (manner of) location verbs and not the (manner of) motion verbs which continue to be used productively in English. In addition, as the OED data as well as those from our corpus show, the posture verbs were still used as locative verbs long after the French influence, whose peak is to be situated in the middle of the 12th century.

The second hypothesis suggested by Lemmens (2014) is that the demise of the posture verbs is due to important changes in the English aspectual system, notably the rise of the *-ing*-form.[7] As a rule, in contemporary English, state verbs

7. This hypothesis was suggested to M. Lemmens by Jaqueline Guéron (personal communication, 20 September 2003).

cannot occur in the *ing*-form, yet the CPVs form a notable exception to this: the CPVs can be used in the *ing*-form to denote a temporary situation (e.g., *She was sitting on the sofa*) while the base form is used for more permanent states (e.g., *The statue stands in the park*). The hypothesis would be that at the time the *ing*-form came into being, its semantics were in conflict with the static semantics of the CPVs, disfavouring their use in locative contexts. The other Germanic languages have *not* developed an equivalent to the *ing*-form and the CPVs cannot occur in the alternative progressive constructions (cf. Lemmens, 2005a, 2015). While this is a possible avenue to be pursued, it remains unclear how to link up the new semantic distinction (temporary vs. permanent state) with the demise of the posture verbs as basic locative verbs. Moreover, an exploratory pilot study (Lemmens, 2005c) on *sit* suggests that the rise of the *ing*-form and decline of posture verbs do not converge.

The third, and possibly most plausible hypothesis, is drawn from Newman (2009, p. 46):

> The English posture verbs, to the extent they develop locative meanings, do so in relatively limited ways and it may be that the fluctuation in action/state interpretations of these verbs is relevant to understanding the limited nature of this type of extension.

In other words, Newman attributes the non-grammaticalisation of English CPVs to a double yet opposite development of these verbs: they come to express the inchoative aspect ***without*** the expected particle expressing the path of motion (***down*** for *sit* and *lie*, *up* for *stand*); at the same time, the verb-particle combinations themselves begin to be used to refer to a static situation (e.g., *sit down for hours*). Also here, English differs from the overall tendency found in the Germanic languages. As pointed out by Talmy (2000) the expression referring to the dynamic interpretation ('getting into a state') is derived from the static expression via the addition of particles (*stand* > *stand* ***up*** ; *sit* > *sit* ***down*** ; *lie* > *lie* ***down***). According to the Oxford English Dictionary, the particle appears in Late Old English but only regularized in Middle English when the phrasal verb system fully established itself (Brinton & Traugott, 2005, p. 123–125). However, in Present-Day English, the use of the particle with CPVs is not as neatly delineated as it seems to be at first sight. Yamagushi & Newman (2002) notice that, for *sit*, it is still possible to have a state interpretation with *down* (which we will call the 'augmented' CPV construction) or a dynamic interpretation without *down* (which we term the 'bare CPV construction'). A similar observation holds for the two other CPVs as well. The differentiation between static and dynamic meanings is clearer in other Germanic languages where there is a sharp differentiation with two distinct lexical items or constructions. In German and Swedish, for example, the intransitive *sitzen/sitta*

'sit' refers to the state sense, while the transitive reflexive *sich (hin)setzen / sätta sig* (litt. 'set oneself (down)') refers to the dynamic sense.[8] Similarly, in Dutch intransitive *zitten* 'sit' refers to the state, *gaan zitten* (litt. 'go sit'), to the inchoative.

It is particularly the development of the dynamic meaning of particle-less *sit*, *lie*, and *stand* that, as Newman has it, has prevented the English CPVs from developing into basic locative verbs. Our corpus has indeed yielded cases of particle-less uses with such dynamic reading (or at least, such a reading is not excluded), as illustrated in the following examples.

(2) a. At the last this poller went somwhat before to knyghtbryge & *sat* upon the brydge (Sit, 1500–1569)
b. The vile wretch prevailed upon Mrs. Bevis to personate me, (…) and persuaded her to *lie* along a couch. (Lie, 1710–1780)
c. "It is a long time since I saw you," he said brokenly. "*Stand* so that I can look at you, Olive!" She obeyed. (Stand, 1850–1920)

While Newman's hypothesis is promising, the empirical study he provides in support is based on only one of the three cardinal posture verbs (*sit*) in a small and fairly non-representative corpus (the King James' Bible). Nevertheless, as Newman's hypothesis seems to be the most promising of the three hypotheses that Lemmens (2014) has suggested, we have taken this one as the starting point of our study which aspires to offer more reliable empirical evidence, based on a careful analysis of the three verbs in a much larger corpus.

3. Corpora and method

Our study analyses the usage of *sit*, *stand* and *lie* in two different corpora covering the period of Modern English. The first corpus is the *Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME) (cf. Kroch et al., 2004). Consisting of 1.8 million words, the corpus covers the period 1500–1710 divided in three 70-year periods. The second corpus is the *Corpus of Late Modern English Texts* (CLMET), version 3.0 (cf. Diller et al., 2011). This corpus contains 34 million words and covers the period from 1710 to 1920, also divided in three 70-year periods: 1710–1780, 1780–1850, and 1850–1920. While different in size, the two corpora are balanced with respect to genre, a compilation of mostly narrative fiction,

8. Notice that in contemporary English *set* is no longer the causative equivalent of *sit*, whereas in German and in Swedish (and partially in Dutch) the intransitive/transitive pairs continue to exist. See Lemmens (2006) or Gullberg (2009) for a more detailed discussion of the causative verbs in Germanic languages.

letters, treatises, and essays. All attestations have been extracted using the free concordance programme *AntConc*;[9] as the corpora are not tagged, our queries have covered via regular expressions all possible inflected forms of the CPVs, i.e. infinitive, third person singular present, simple past, and present and past participle as well as all variant forms of these attested in the OED for Early Modern English, such as *lyeth, ligge, sitten, syteth, stoud, standith*. In total, for the three CPVs, there were 3,248 attestations in the PPCEME and 47,675 attestations in the CLMET. Given the small size of the PPCEME, all the attestations from this corpus have been coded for our analysis. However, to keep matters manageable, we have restricted our detailed analysis for the Late Modern English corpus (CLMET) to a random selection of 500 occurrences for each verb for each period.[10] This gives a total of 1,500 sentences for each verb for the whole Late Modern English period and a grand total of 4,500 sentences.

Some of the attestations extracted via AntConc were incorrect and did not concern CPVs; these were all manually removed.[11] For the CLMET, each such incorrect attestation deleted from the random selection was replaced by the next attestation in the random list for the verb in question, such that the total of 500 attestations per verb per period was preserved. Once all the noise was removed, we had for the two corpora and for the six periods 1,895 attestations for *sit*, 2,603 for *stand*, and 2,405 for *lie*. These attestations were subsequently annotated in a spreadsheet according to eleven different parameters, the most important of which are the type of Figure (human, inanimate, concrete or abstract), the usage of the verb (postural, locational, metaphorical, copular), the presence of a Ground specification (for location events), the presence of a nominal or adjectival complement and, most importantly for testing Newman's hypothesis, the presence of the particle (for the postural uses only).

The different parameters allow us to take a more constructional view on the uses of the CPVs focusing on three constructions: (i) the postural construction, (ii) the locational construction and (iii) the copular construction as presented in Section 1.

9. http://www.laurenceanthony.net/software/antconc/

10. The random selection was made by assigning a random number to each sentence (using the Random function in Excel) and then sorting the sentences in decreasing order; the first 500 were subsequently selected for further analysis.

11. These inaccuracies concerned, for example, homographic forms, such as the Latin word *sit* (inflected form of *esse* 'be'). Also, for *lie*, there is the confusion with the agentive form *lay* which is homophonous with the past tense form of *lie*, with the noun *lie* meaning 'falsehood', and with the corresponding verb *lie* ('not tell the truth'). Similarly, the regular expressions (in the query for *stand*) also retrieved attestations with *staid* (= stayed).

Unfortunately, for the period of Early Modern English within the PPCEME, the attestations with a postural use of the CPVs were too small to test Newman's hypothesis. For this particular part of the study (presented in Section 4.3 below), an additional corpus was used for this period, the Early English Books Online corpus (EEBO),[12] from which 756 occurrences with postural meaning were coded for *sit*, 294 for *stand*, and 206 for *lie*. This extra corpus was only used for the testing of Newman's hypothesis.

4. Results

The two main research questions that our corpus study addresses are: (i) What was the degree of grammaticalisation of the English CPVs, especially in the earlier periods? and (ii) Can we find evidence for the demise of the grammaticalised uses of the English CPVs, notably when and why it occurred? This section presents the results of our corpus analysis in answer to these questions. In the first subsection (4.1), we look at the locative uses of CPVs over the different periods, to establish their status as basic locative verbs and how those uses have evolved over time.[13] The more revealing case study, however, is that of copular constructions as the most grammaticalised uses of CPVs (Section 4.2) which we will analyse from both a quantitative and qualitative perspective. The last Section (4.3) will empirically evaluate Newman's hypothesis for the demise of CPVs in English, by aligning the frequency of the dynamic uses of bare CPV constructions with that of the grammaticalised copular uses.

4.1 Locational uses

While Present-Day English tends not to use CPVs to express the location of entities in space but prefers the neutral verb *be*, English used to employ CPVs to express static location events much more frequently in earlier periods. This clearly suggests that there must have been a progressive decline of such locational uses of CPVs; as such, they can be said to have (unexpectedly) missed further entrenchment as basic locative verbs, also blocking any further grammaticalisation.

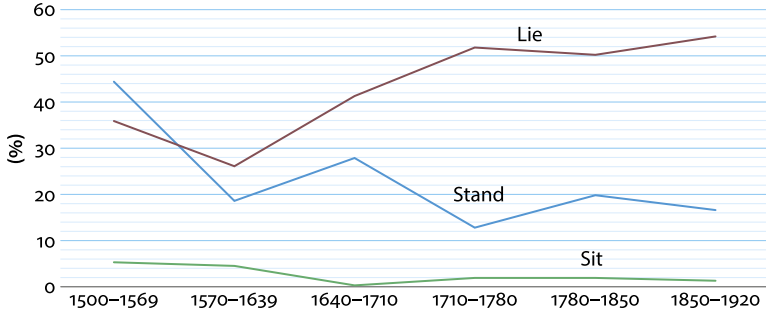
12. Developed by Petré (2014), the Early English Books Online corpus counts 525 million words covering the period 1474–1700. Though the corpus has its own concordancer, AntConc was used as for the other corpora.

13. As said, Kuteva (1999) considers the use of CPVs as basic locative verbs a precondition for their further grammaticalisation to, e.g., aspectual markers.

Tables 1 and 2 and Graph 1 present the frequency of the locative uses of CPVs over the different periods covered by our corpora (the graph gives the evolution in relative frequencies with relation to the other uses only, Table 2 sets off the frequencies for the locational uses from all other uses).

Table 1. Locative uses of CPV: Raw frequencies per period

	1500–1569	1570–1639	1640–1710	1710–1780	1780–1850	1850–1920
sit	8	5	0	8	8	5
stand	203	77	65	64	99	83
lie	131	81	95	259	251	271
Total	**342**	**163**	**160**	**331**	**358**	**359**



Graph 1. Evolution of the locational construction

Table 2. Evolution of the locational construction vis a vis other constructions (expected frequencies between brackets)

		1500–1569	1570–1639	1640–1710	1710–1780	1780–1850	1850–1920
sit	locational construction	8 (*2.87*)	5 (*2.14*)	0 (*2.08*)	8 (*8.97*)	8 (*8.97*)	5 (*8.97*)
	other	152 (*157.13*)	114 (*116.86*)	116 (*113.92*)	492 (*491.03*)	492 (*491.03*)	495 (*491.03*)
		$\chi^2 = 17.37$; $df = 5$; $p = 0.0038$					
stand	locational construction	203 (*103.76*)	77 (*93.77*)	65 (*52.9*)	64 (*113.52*)	99 (*113.52*)	83 (*113.52*)
	other	254 (*353.24*)	336 (*319.23*)	168 (*180.1*)	436 (*386.48*)	401 (*386.48*)	417 (*386.48*)
		$\chi^2 = 171.22$; $df = 5$; $p < 0.00001$					
lie	locational construction	131 (*168.56*)	81 (*139.72*)	95 (*103.66*)	259 (*225.35*)	251 (*225.35*)	271 (*225.35*)
	other	243 (*205.44*)	229 (*170.28*)	135 (*126.34*)	241 (*274.65*)	249 (*274.65*)	229 (*274.65*)
		$\chi^2 = 92.77$; $df = 5$; $p < 0.00001$					

As these data show, *sit* has always been very restricted in its use in the locational construction compared to the two other CPVs; nevertheless, its decline is noticeable from 1500 down to 1920 with a sharper decrease in the 17th century. A similar trend, but more pronounced, has occurred for *stand*; while overall the evolution has a clear falling contour there are two small peaks for this verb, one in the second half of the 17th century and one about a century later. Inspection of the data does not reveal any particular reason for these peaks (such as possibly high frequency of particular idiomatic uses), which means we cannot exclude it being an effect of the data sampling.

A chi-square test reveals that from the first period to the third, the frequency of locational constructions is higher than expected. For the later periods, their frequency decreases in comparison to what is expected, except for *lie* for which, quite strikingly, the locational construction markedly increases over time. A closer look at the data reveals that *lie* got entrenched in its metaphorical locational use with abstract Figures or Grounds. Allowing a wide range of abstract inanimate entities, *lie* is in fact the most metaphorical of the three verbs (as is also the case in Dutch). It has continued such metaphorical uses up till the present day, such as in:

(3) The difficulty *lies*, as it always does, on the border lines between the respective spheres of influence: Logic and Faith. (Lie, 1850–1920)

(4) Those were the men whom, if the choice had *lain* with him, he would have wished to command. (Lie, 1850–1920)

While some data sampling effects cannot be fully excluded for the intermediate periods, the data do reveal that the locational uses were fairly entrenched in the earlier periods, but that they statistically significantly declined in the evolution to Present-Day English, with metaphorical uses of *lie* being a clear exception to the overall trend.

4.2 Copular constructions

4.2.1 *Quantitative analysis*

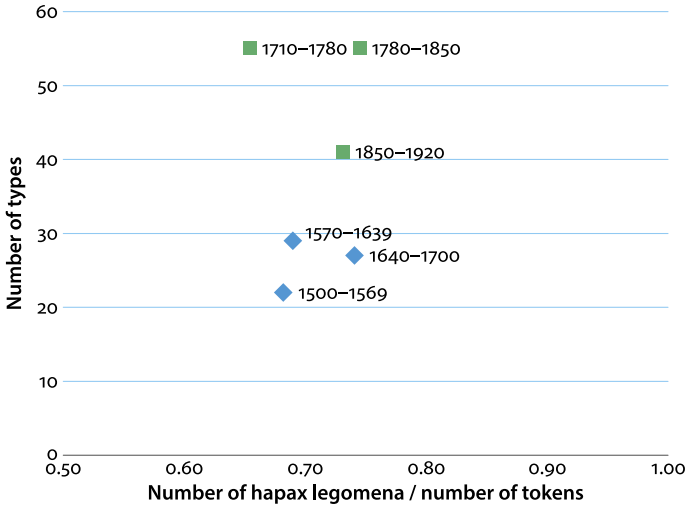
Following the assumptions of Construction Grammar, we can measure the degree of entrenchment of a certain construction by its degree of productivity. Under this view, we can establish the degree of grammaticalisation of CPVs by measuring to what extent they could be used productively in the Modern English period in the most grammaticalised construction, the copular construction. A reliable method for measuring productivity that has been suggested by Baayen & Lieber (1991) is to divide the number of *hapax legomena* (types that occur only once in the sample) by the total number of tokens; the underlying assumption is that hapaxes can

be taken as indications of the degree to which a construction is open to new uses ('one off' cases). In other words, the higher the number of hapaxes compared to the total number of tokens, the higher the productivity rate. However, this productivity rate is not sufficient since in our diachronic perspective, we also want to measure the increase or decrease of ***types*** (in our case, that of the Subject Complement Phrase) over the different periods. The more types a construction has, the more productive it is, and the more entrenched it may be. In addition, the variation in types also indicates the degree of host-class expansion. Following the method suggested by Petré (2012) on the English *wax* copula, the productivity of the construction can thus be represented in a graph which combines these two productivity indexes (which gives what Baayen (1992) has called the *global productivity*): the Y-axis presents the number of types, the X-axis, the above mentioned productivity rate (number of hapaxes divided by the total number tokens, which gives values between 0 and 1).[14] The theoretical view adopted here is quite compatible with Bybee's exemplar theory (see Bybee, 2006, 2010) which also looks at category expansion and exemplar clustering.

The following graphs show the productivity for each CPV in the adjectival copular construction; a similar analysis of the nominal construction was not possible because of too low frequencies. A cluster of points in the top right-hand corner reflects high productivity as it implies a high number of types as well as a high productivity rate, a cluster in the top left-hand corner reveals less productivity with a high number of types but with all types occurring only once. In fact, if the X value equals 1, this means that the construction's types are all *hapaxes*. Turning to the results for the CPVs, we can see that the values for *sit* are closer to 1 than those for *stand* and *lie*. The Y-axis shows that as for the number of types, *stand* and *lie* are much higher, whereas *sit* shows fewer types.

Analysing the evolution of the productivity index for each CPV in the graphs is to be done carefully since the two corpora, 1500–1710 and 1710–1920, are not of the same size (but the time periods within the corpora are). This is reflected in our graphs by two subclusters: one for Early Modern English, logically lower in types because of the reduced size of the corpus, and one for Late Modern English, higher on the Y axis because of the larger size of the corpus. Therefore, the two corpora are to be interpreted independently with respect to the Y-axis, but can be interpreted together for the X-axis whose value has been rescaled to a range between 0.5 and 1.[15]

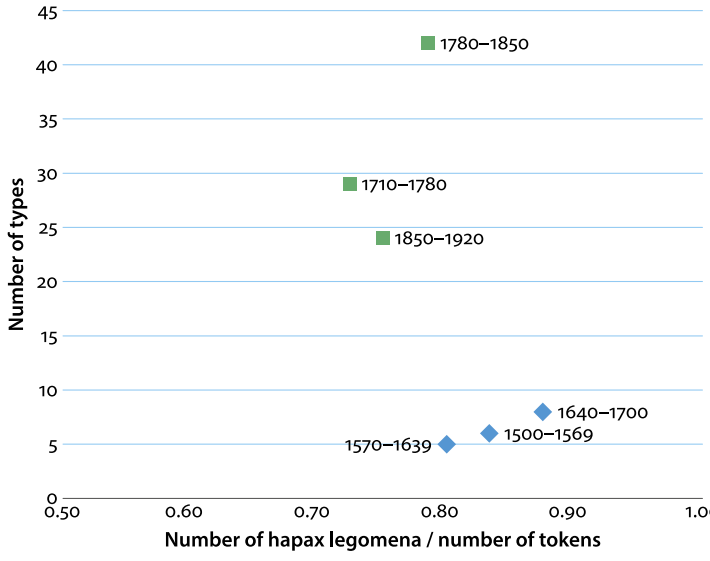
14. See also Schröder (2011, pp. 51ff) and Barðdal (2008, pp. 24ff) on these measurements of productivity and for some discussion on productivity in general. On the latter, see also, among others, Plag (1999) and Dal & Namer (2015).



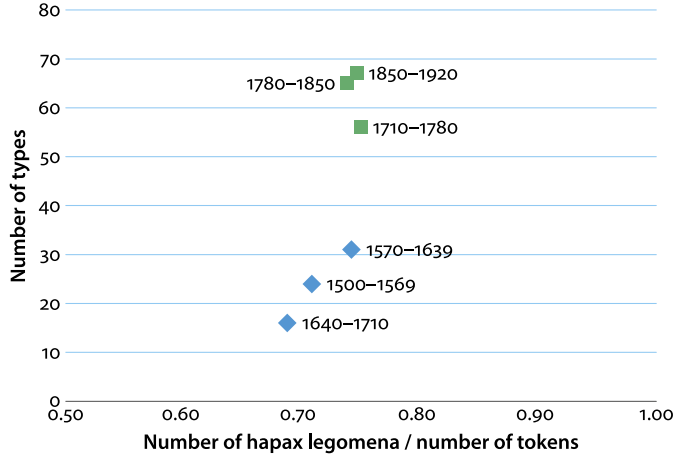
Graph 2. Global productivity for [SubjectNP + *stand* + SubjCompAdjP]

A number of conclusions can be drawn from these calculations. Firstly, for *stand*, both in Early and Late Modern English, the number of types decreases and so does the productivity rate over the two corpora. Such evolution suggests a fixation or consolidation of some collocations; this will indeed be confirmed below by the qualitative analysis. Secondly, for *sit*, in Early Modern English, the number of types increases and the productivity rate increases as well. In Late Modern English, there are fewer types but the productivity rate increases. Yet, globally for both corpora, the productivity rate (i.e., the ratio between the number of hapax legomenon and the number of tokens) decreases. This confirms the same tendency as for *stand*, *viz.* that some adjectives might form fixed collocations with the CPV and remain used throughout the years. Finally, for *lie*, in contrast, the X-value increases across the two corpora; while the number of types decreases in Early Modern English, it slightly increases in Late Modern English. This suggests that *lie* allows a wide range of adjectives but that they are not recurring enough to form fixed expressions. All these hypotheses need to be confirmed by a close qualitative analysis of the copular constructions, given next.

15. As Baayen (2008, pp. 223ff) points out, such productivity measures are sensitive to corpus size: as corpus size increases the chance of finding a new hapax legomenon will decrease. However, this is not an obstacle here, as the corpus sizes are not all that different; for Late Modern English, we used a random sample of 500 and the trend is clear. In addition, the data are not from one and the same text (unlike as in Baayen's case study) but are drawn from randomly selected texts that are fairly independent of each other. More important is that the quantitative analysis does confirm our qualitative analysis.



Graph 3. Global productivity for [SubjectNP + *sit* + SubjCompAdjP]



Graph 4. Global productivity for [SubjectNP + *lie* + SubjCompAdjP]

4.2.2 *Qualitative analysis*

For the adjectival copular construction, all the adjectives have been categorized based on their semantics. There are three categories: (i) locational adjectives (e.g., *deep, high*, *near*, etc.) specifying the location in space of a Figure; this category also contains relator adjectives (e.g., *close to, next to*, etc.); (ii) adjectives pertaining to motion or the absence of it, which brings about a postural view of the event (e.g., *tight, firm, still* when in post verbal position, cf. Maizak, 2002); (iii) all the

other adjectives. This article focuses mainly on the third category of adjectives, since being least tied to postural or locative semantics, they are most interesting for measuring the degree of grammaticalisation. Except for *sit*, this category is the largest one, gathering more than half of all the adjectives that occur with CPVs A close analysis of collocates reveals that there are indeed clusters of use directly linked to what Newman (2002) has called the experiential clusters associated with each of the verbs, as further illustrated below.[16]

Right from Early Modern English on, *lie* co-occurs with adjectives referring to WEAKNESS such as *oppressed, sick, weak, prostrate, bedrid, concealed*, or *helpless.* There is also a cluster around LETHAL HARM or DEATH, e.g., *burried, dead, breathless, bloody*, *wonded* ('wounded'). Another cluster is linked to INACTIVITY with adjectives such as *asleep, idle, dormant, flat, becalmed, unactive*, with *asleep* and *idle* being the recurring collocates. There are also some other fixed expressions in our corpora such as *lie open* with inanimate Figures as in *The Oracles of God lie open to us* (Lie, 1640–1710, 116). This construction is also found with *stand*.

For *stand*, there is also a whole cluster of adjectives directly linked with the concepts of POWER and DOMINANCE that are frequent right from the Early Modern English period as well, e.g. (categories (ii) and (iii) conflated for convenience) *victorious, inflexible, rooted, invested, infaillible* (sic.), *secure, armed, proudliest*, or *firm*. *Stand* is also associated with RESISTANCE as reflected in our corpus with the recurring construction *stand fixed in* X, with X being *thought*, *silence*, or *meditation.* This idiomatic expression combines the notion of RESISTANCE with that of INACTIVITY of the X slot components usually not associated with a standing position.

For *sit*, two adjectives that appear several times in the adjective slot in Late Modern English are *pensive* and *attentive* which can be linked to the idea of INTELLECTUAL ACTIVITY that is typically associated with this verb (cf. Newman, 2002). CONTAINMENT-*sit* (cf. above) is also found in adjectives such as *absorbed in reflection, shrouded*, or *framed.* Surprisingly, from the Early Modern English period onwards, *sit* also collocates with adjectives that one would a priori associate with *lie* and that express INACTIVITY and WEAKNESS like *idle, sad, breathless, composed, abashed, confounded, desolate*, or *disordered*; similar ones have also been attested with *stand* (*fearful, indicted, repealed, sick, aghast, astonished, confused*, and *defenceless*). This shows that the three CPVs share the semantic domains of WEAKNESS and INACTIVITY while intuitively one may not have expected this. In

16. Experiential clusters are bundles of characteristics that are associated with the sitting, standing and lying postures, and pertain to specifications in the force-dynamic domain (pertaining to balance, physical effort) or in the social domain (activities commonly associated with these postures), etc. See Newman (2002) or Lemmens (2002) for more on this.

fact, there is a whole set of adjectives for each CPV which do not share any semantic features (or very few) but which can collocate with two or three CPVs, as is the case with motion and location adjectives. For instance, *sit/stand/lie alone* is a construction which seems to be productive for each of the three CPVs. It even becomes more and more common for *sit* and *stand* in the third period. Another noticeable fixed expression is *sit/stand silent.* While this construction was more frequent with *sit* in the first period, its frequency decreases in the second and third period and its occurrence with *stand* increases along the years.

The semantic domains shared between the three CPVs could be schematized as follows:

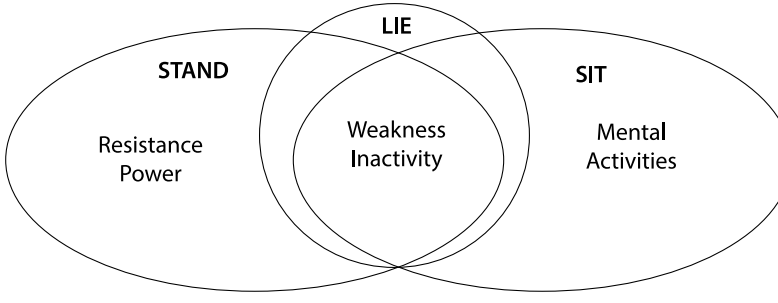


Figure 4. Semantic domains for the adjectival copular construction

The fact that they share these adjectives, which at first sight may not be in line with the 'usual' associations of the CPV in question, could be taken as evidence of their highly bleached character in these adjectival copula constructions where they have become almost interchangeable copulas.

In contrast to the adjectival construction, the nominal copular construction is not only much less frequent but the collocates of the three verbs are also much more diverse. There are nonetheless important tendencies to be observed with respect to the noun collocates occurring in this construction. First of all, *lie* stands out in this respect since there is no recurring pattern (i.e., semantic clustering) for the nominal construction. For *sit*, the collocating nouns all refer to 'sitting professions'. Some nouns are introduced by the preposition *as* (prepositional copular construction) as in (5):

(5) Troy saw his unconscious wife on high before him as described, *sitting* **as** queen of the tournament. (Sit, 1850–1920)

In our corpora, the construction *sit as* + *Noun* is attested with the following nouns: *chefest* ('chief'), *criminal judge, God, judge, guardian, member, queen, stuart.* The nouns can also be directly linked to the CPV without the preposition, as it is the case with *bishop, supervisor, state auctioneer*, and *umpire*. Even though all

these nouns appear only once in the corpora (except for *sit judge* which is a recurring collocation), the tendency for *sit* to occur with 'sitting professions' is clear.

Finally, for *stand*, there are different semantic clusters to be noticed. First, there is a set of 'standing professions': *knight, barrister, souveraign, spectator, arbitrator, bottle-holder, trumpeter, guard, member* and, most frequent, *sentry|sentinel|centinel*. The common point to these nouns is their referent's standing position while at work. A standing position also motivates the phrase *stand as an accused person* where the person can easily be depicted as standing in the witness box. Another semantic domain expressed by the collocating nouns is that of REPRESENTATION as in *stand as embodiment* or *stand proxy*. This notion of REPRESENTATION is also visible with the (recurring) noun collocate *godfather* as in (6):

(6) [He was] condescended to toy with my Little Ones; appeared to take a huge Liking to my two eldest Boys; *stood* Godfather to my little Girl that is now in her Mother's Arms. (Stand, 1710–1780)

In this example, the collocation of *stand* with *godfather* can be interpreted in two ways: either one claims that the person is standing for the father of the little girl in front of God – a religious interpretation that hinges on the image of REPRESENTATION, or one considers FAMILY to imply SUPPORT, a notion that *stand* is also associated with (cf. Gibbs et al., 1994). In fact, another frequent collocate, *friend* as in (7), points towards this second interpretation:

(7) And though wee had a Caffeloe or Conuoy of two hundred strong, yet wee were glad to intreate the Captaine of that vnruly Crew to *stand* our friend. (Stand, 1570–1639, 243)

There is no overt link between being someone's godfather or friend and the standing position, but there is certainly a metaphorical link whereby one's friends and godfathers are people one trusts and can rely on; they are thus standing next to you and take responsibility of your well-being.[17] Clearly, within the nominal and prepositional copular constructions, the CPVs demonstrate clearer distinctive semantic domains of collocates.

Most importantly, this section proves that while English CPVs could be used as copulas, they did so in a very constrained way; these uses were restricted to a number of relatively fixed expressions that can mostly still be related to the CPVs' experiential clusters. This is especially true for the nominal constructions, whose

17. One of the anonymous reviewers points out that there still is a link with a standing posture, as the godfather would be standing next to the baptismal font. This may indeed be plausible for stand godfather, but not for the collocation with friend, where the support reading seems the only one applicable.

anthropocentric meaning has not been bleached completely; for the adjectival copular construction, there seemed to have been a higher degree of bleaching, almost to the point of the verbs having become interchangeable for some adjectives. In other words, the English CPVs did show patterns of (budding) grammaticalisation, but they may not have reached a full grammaticalised status. Why this is so will be discussed next.

4.3 Putting Newman's hypothesis to the test

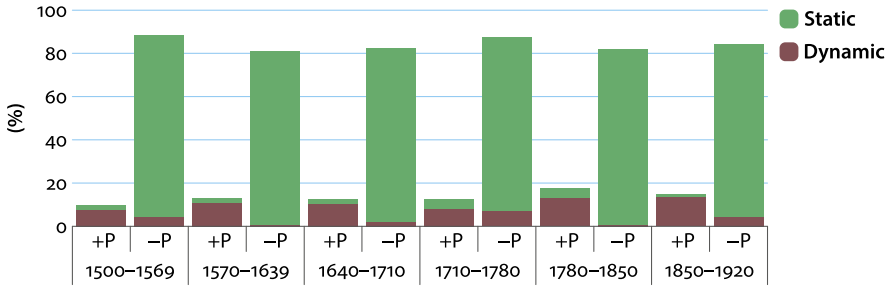
As will be recalled, Newman (2009) suggested that the evolution of the English CPVs to basic locative verbs and thus also their further grammaticalisation was "compromised", in Newman's (2009) terms, by the rise of the use of bare CPVs (i.e., without a particle) to refer to the dynamic event of getting into a standing, sitting or lying position.

In order to test Newman's hypothesis, we need to look at the postural construction; more specifically, we need to compare the constructions with and without particle and what type of event they refer to. Recall that neither of these constructions can be exclusively restricted to the static or the dynamic interpretation, Newman's hypothesis being that it is precisely this fuzziness in the system that prevented the English CPVs to further grammaticalise into locative constructions. As mentioned earlier, because of the low number of postural uses for the period of Early Modern English, we have augmented our corpus with postural uses from the EEBO corpus (756 for *sit*, 294 for *stand*, and 206 for *lie*) so as to have a more reliable sample.

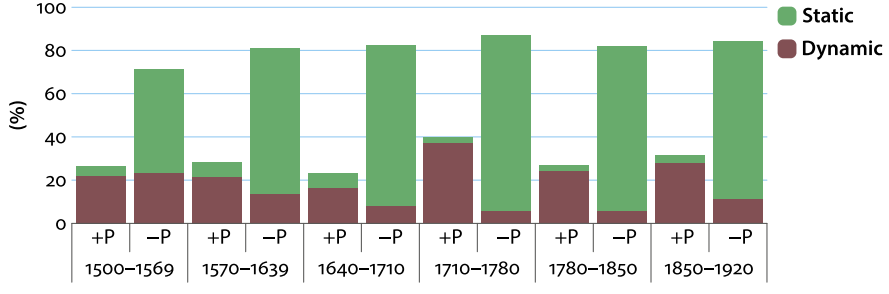
Our analysis relies on clustered and stacked columns allowing us to trace the evolution of the dynamic and static postural uses and how they are coded, without a particle (−P) or with a particle (+P); via this method, we also exclude that the increase of the bare dynamic use is due to the increase of the dynamic uses overall.[18]

As these graphs show, the static use of CPVs dominates overall, confirming Dunn's et al. (2007) findings, except for *sit* which appears much more divided and does not favour the static interpretation, as also stated in Newman (2009). Furthermore, on the whole, all throughout the Modern English period, the expected prototypical constructional patterns have been attested: the bare construction, i.e. *Subject + CPV (+ preposition + Ground)*, refers to a static event while the augmented construction *Subject + CPV + particle (+ preposition + Ground)* expresses a dynamic event. Most importantly though, the static-dynamic fuzziness mentioned by Newman is also visible in these graphs, as there are clearly cases of

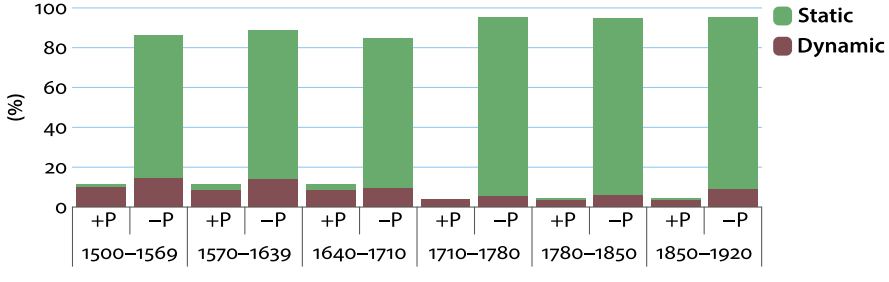
18. In the graphs, the bars for +P and −P taken together represent 100%.



Graph 5. Evolution of the static and dynamic interpretations within particle and particle-less constructions for *lie*



Graph 6. Evolution of the static and dynamic interpretations within particle and particle-less constructions for *sit*



Graph 7. Evolution of the static and dynamic interpretations within particle and particle-less constructions for *stand*

particle constructions with a static interpretation and cases of particle-less constructions with a dynamic interpretation.

However, the evolution of the construction *Subject + CPV (+ preposition + Ground)* with a dynamic interpretation is not clear. Some increase is noticeable

for *sit* and *stand* in Late Modern English, from 1710, but not for Early Modern English, the period on which Newman bases his hypothesis. Moreover, when conflating the evolution of the particle-less construction with the evolution of the dynamic use as a whole, it appears that both are decreasing for all the three CPVs, except for the last two periods 1780–1850 and 1850–1920, where both the dynamic interpretation and the particle-less dynamic construction are increasing. Therefore, while there is no indication of the non-prototypical construction gaining ground on the postural uses, there is a slight increase of the construction but not for the period Newman claims.

Evaluating Newman's hypothesis against corpus data also raises another issue. As the following table reveals, we are looking at a very specific construction, with fairly low frequencies. (The ambiguous cases are those where the context allowed both interpretations.)

Table 3. Particle and particle-less constructions in Modern English

			PPCEME + EEBO			CLMET		
Verb	**Interpretation**	**Particle**	**1500–1569**	**1570–1639**	**1640–1710**	**1710–1780**	**1780–1850**	**1850–1920**
lie	**Dynamic**	+p	14	16	10	7	22	22
		−p	8	1	2	6	1	7
	Static	+p	4	5	2	4	7	2
		−p	158	123	78	70	134	127
	Ambiguous	+p	1	1	0	0	0	0
		−p	3	2	5	0	1	1
sit	**Dynamic**	+p	71	67	41	142	93	113
		−p	74	42	20	20	20	44
	Static	+p	14	6	18	9	11	13
		−p	158	198	169	211	271	240
	Ambiguous	+p	3	2	4	0	0	0
		−p	8	7	10	3	2	0
stand	**Dynamic**	+p	28	19	10	7	8	10
		−p	42	31	11	10	14	26
	Static	+p	4	4	3	0	2	2
		−p	200	155	84	154	202	240
	Ambiguous	+p	0	0	0	0	0	0
		−p	5	6	4	0	0	0

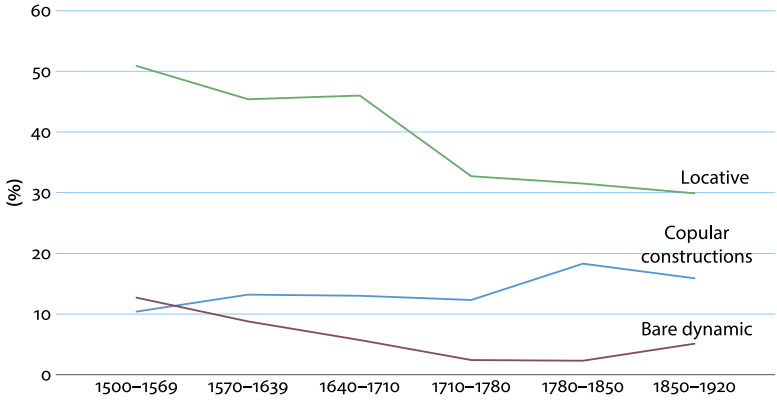
This table does show the fluctuation in the system that Newman has pointed out. Notably, *stand* occurs more in the construction without the particle for the

dynamic interpretation. The rows with a bold frame are the non-prototypical constructions: the dynamic in a particle-less construction, and the static with a particle. These constructions are, according to Newman (2009), the ones that are to be held responsible for the fuzziness in the English CPVs system. However, these non-prototypical constructions represent a modest portion of the whole set of occurrences, invariably under 4% of the total number of occurrences per period.[19] This may compromise the validity of Newman's hypothesis: to what extent can a phenomenon restricted to such a modest set tilt the whole system of English CPVs preventing them from further grammaticalisation? At first sight, it seems unlikely that it would have disrupted the overall grammaticalisation of CPVs. However, these postural uses do constitute the prototypical uses of the CPVs, and prototypical centres are highly influential for the evolution of the entire category. Thus, the impact of the fluctuation between the static and dynamic interpretations may have been much higher than what the lower percentages suggest, since the development of CPVs into locative markers exclusively depends on the prototypical anthropocentric uses (Dunn et al., 2007). In sum, Newman's hypothesis may still be correct but our more extensive corpus provides no firm statistical evidence to confirm it.

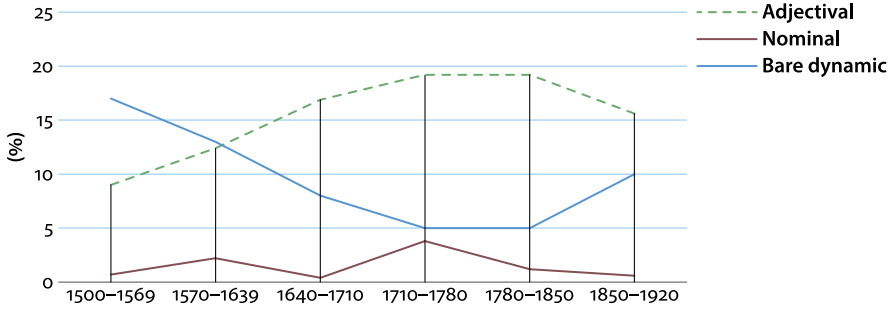
While the percentages of bare dynamic uses (i.e., without a particle) by themselves may not give a clear picture, our data allows to set them off against the frequency of locative and copular uses. In other words, if Newman's hypothesis is correct that the rise of the bare dynamic postural use is responsible for the demise of the English CPVs, then this evolution, even if minimal, should lead to a decrease of the locative and copular uses. This is what is presented in Graph 8, from which one can also infer that the postural uses dramatically increase in relation to the other constructions, a general indication of the halted grammaticalisation of the CPVs already shown above.

The picture that emerges is not entirely clear. First of all, the bare dynamic postural uses first go down, and then up again, but the percentages are very low and fairly close. The locative uses continue to decrease over the entire period. However, what is perhaps most relevant here is to look at the copular constructions, as the most grammaticalised uses of the CPVs. Their uses do rise at first, but decrease later on which co-occurs with the (slight) increase of the bare dynamic construction occurring between 1780 and 1920. The situation becomes clearer when we look at the copular uses for each of the CPVs individually, as presented in the graphs 9, 10, and 11. The evolutions of copular constructions

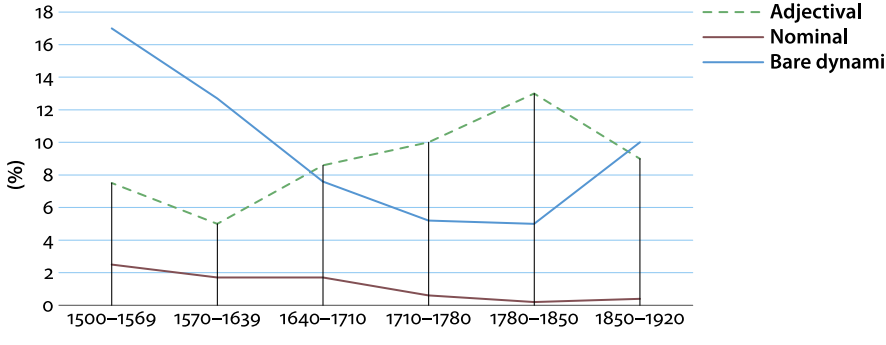
19. Only *sit* shows a larger portion for the bare dynamic construction with 11.2% in the first period, and 6.8% in the second period.



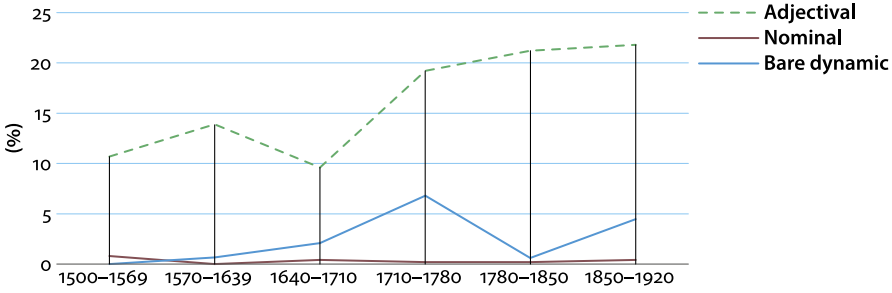
Graph 8. Evolution of the bare dynamic postural, locative and copular constructions (vis a vis all uses)



Graph 9. Conflating the evolutions of the bare dynamic construction with the copular constructions for *stand*



Graph 10. Conflating the evolutions of the bare dynamic construction with the copular constructions for *sit*



Graph 11. Conflating the evolutions of the bare dynamic construction with the copular constructions for *lie*

with *sit* and *stand* are quite clear: their decrease co-occurs with the rise of the bare dynamic uses.

In contrast, *lie* does not display any such inverse evolution (the bare dynamic construction increasing and the copular constructions decreasing) in Late Modern English. In fact, this evolution occurs earlier between the bare dynamic construction and the adjectival construction, in Early Modern English between the second and third periods. *Lie* has always behaved differently from its two companions *sit* and *stand*, notably because of its wide metaphorical side.

In sum, our data confirms some confusion between static and dynamic uses and construction type (with or without particle); however, the rise of the bare dynamic use is only minimal and while it does co-occur with an overall decrease of CPV use, it still remains an open question whether this rise is indeed the (sole) triggering factor.

5. Conclusion

Our corpus analysis reveals evidence that the English CPVs clearly had the potential to grammaticalise further into locative markers and copulas as they did in other languages. Our data reveal that they did grammaticalise to some extent, these verbs being amply used in locative constructions and copular constructions. For the latter type, while the adjectival copular construction was clearly more entrenched than the nominal one, the three CPVs display what could be referred to as 'grammaticalised islands' with very fixed collocations which persisted throughout the period of Modern English, but eventually disappeared.[20] Our data

20. Our use of the term 'island' is not to be confused with 'niches' (cf. Torres Cacoullos & Walker, 2009) which retain a clear link with the original meaning and usage which is not the case here.

only partially confirm Newman's (2009) hypothesis attributing this demise to the constructional fuzziness between the dynamic and the static interpretations of the CPVs. Qualitatively, there is indeed some fuzziness in the system with the construction without the particle being interpreted as dynamic, yet quantitatively no clear and firm evolution of this construction could be found. Moreover, this hypothesis concerns a limited phenomenon as it only applies to a tiny proportion of the whole data set (postural uses). However, as this concerns the CPVs' prototypical centre, it could still be maintained that it affected the entire CPV system in English. Zooming in on the copular uses, our analysis indeed reveals that the increase of the bare dynamic construction co-occurs with the decrease of the copular constructions. This inverse movement is very clear in Late Modern English for *sit* and *stand* but not for *lie*. The explanation for this may be found in the higher entrenchment of metaphorical uses of this verb, especially in the context of the location of abstract entities.

As an overall conclusion, our analysis confirms Traugott's view that "each micro-construction has its own history within the constraints of larger patterns" (2014, p. 17). Looking at the constructions in which the English CPVs have occurred has allowed a more accurate study of their degree of grammaticalisation in earlier periods, and a more careful verification of the causes of their demise. One of the aims of our paper has been to evaluate the most plausible hypothesis that has been suggested in the literature for this demise, i.e. the rise of the bare dynamic construction (Newman, 2009). The picture that has emerged is a nuanced and complex one, where certain constructions linger on as grammaticalised islands. Further research is necessary to confirm with more reliability what the exact causes were, if any can be found, for the compromised grammaticalisation of the English cardinal posture verbs.

References

Baayen, H. (2008). *Analyzing Linguistic data. A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Baayen, H. R. (1992). On Frequency, transparency, and productivity. In G. E. Booij, & J. Van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 181–208). Dordrecht: Kluwer Academic Publishers.

Baayen, H. R., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, 29, 801–843. https://doi.org/10.1515/ling.1991.29.5.801

Barðdal, J. (2008). *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8

Brinton, L. J., & Traugott, E. C. (2005). *Lexicalization and Language Change: Research Surveys in Linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615962

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

Bybee, J. (2010). *Language, Usage and Cognition*. Oxford: Oxford University Press. https://doi.org/10.1017/CBO9780511750526

Coussé, E. (2018). Grammaticalization, host-class expansion and category change. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category Change from a Constructional Perspective*. Amsterdam: John Benjamins.

Dal, G., & Namer, F. (2015). La fréquence en morphologie: pour quels usages? *Langages*, 197, 47–68. https://doi.org/10.3917/lang.197.0047

Diller, H., De Smet, H., & Tyrkkö, J. (2011). A European database of descriptors of English electronic texts. *The European English Messenger*, 19, 21–35.

Dunn, M., Margetts, A., Meira, S., & Terrill, A. (2007). Four languages from the lower end of the typology of locative predication. *Linguistics*, 45, 873–892. https://doi.org/10.1515/LING.2007.026

Gibbs, R., Beitel, D., Harrington, M., & Sanders, P. (1994). Taking a stand on the meanings of *stand*: Bodily experience as motivation for polysemy. *Journal of Semantics*, 11, 231–251. https://doi.org/10.1093/jos/11.4.231

Gullberg, M. (2009). Reconstructing verb meaning in a second language how English speakers of L2 Dutch talk and gesture about placement. *Annual Review of Cognitive Linguistics*, 7, 221–244.

Himmelmann, N. P. (2004). Lexicalization and grammaticization. Opposite or orthogonal? In W. Bisang, N. P. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization? A look from its components and its fringes* (pp. 21–42). Berlin: Mouton de Gruyter.

Kroch, A., Santorini, B., & Delfs, L. (2004). *Penn-Helsinki Parsed Corpus of Early Modern English*.

Kuteva, T. A. (1999). On 'sit'/'stand'/'lie' auxiliation. *Linguistics*, 37, 191–213. https://doi.org/10.1515/ling.37.2.191

Kvist Darnell, U. (2008). *Pseudosamordningar i svenska, särskilt sådana med verben sitta, ligga och stå*. PhD dissertation. Stockholm: Stockholm University.

Lemmens, M. (2002). The semantic network of Dutch posture verbs. In J. Newman (Ed.), *The Linguistics of Sitting, Standing, and Lying* (pp. 103–139). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.51.07lem

Lemmens, M. (2005a). Aspectual posture verb constructions in Dutch. *Journal of Germanic Linguistics*, 17, 183–217. https://doi.org/10.1017/S1470542705000073

Lemmens, M. (2005b). Motion and location: toward a cognitive typology. In G. Girard-Gillet (Ed.), *Parcours linguistiques. Domaine anglais* (pp. 223–244). Publications de l'Université St Etienne.

Lemmens, M. (2005c). Aspectually induced lexical loss in English. In First International AFLiCo Conference From gram to mind: grammar as cognition. Université de Bordeaux 3, France, 20–22 May.

Lemmens, M. (2006). Caused posture: experiential patterns emerging from corpus research. In Stefanowitsch, A. & Gries, S. (Eds.) Corpora in Cognitive Linguistics. *Corpus-Based Approaches to Syntax and Lexis* (pp. 261–296). Berlin: Mouton de Gruyter.

Lemmens, M. (2017). A Tale of Two Progressives. In A. Makarova, Dickey, S. M. & Divjak, D. S. (Eds.) *Thoughts on Language: Studies in Cognitive Linguistics in Honor of Laura A. Janda* (pp. 175–194). Bloomington, IN: Slavica Publishers.

Lemmens, M., & Perrez, J. (2010). On the use of posture verbs by French-speaking learners of Dutch: a corpus-based study. *Cognitive Linguistics* 21, 315–347.

Lemmens, M. (2014). Une grammaticalisation ratée? Une étude diachronique de stand en anglais. *Anglophonia*, 18. http://anglophonia.revues.org/327; 10.4000/anglophonia.327.

Lemmens, M. (2015). Zit je te denken of ben je aan het piekeren? *Persistentie in het synchrone gebruik van de PREP- en POS-progressiefconstructies in het Nederlands* [*Persistence in the synchronic use of PREP- and POS progressives in Dutch*]. *Nederlandse Taalkunde*, 20, 5–36.

Maisak, T. A. (2002). *Tipologija grammatikalizacii konstrukcij s glagolami dvizhenija i glagolami pozicii* [*Grammaticalization paths of motion and posture verbs*]. Philological Faculty Dissertation, Completed Degree, Moscow State University.

Newman, J. (2002). A cross-linguistic overview of Posture verbs 'sit', 'stand' and 'lie' In J. Newman (Ed.), *The linguistics of sitting, standing and lying* (pp. 1–24). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.51.02new

Newman, J. (2009). English posture verbs: an experientially grounded approach. *Annual Review of Cognitive Linguistics*, 7, 30–57. https://doi.org/10.1075/arcl.7.02new

Newman, J., & Rice, S. (2006). Patterns of usage for English *sit* , *stand*, and *lie*: A cognitively-inspired exploration in corpus linguistics. *Cognitive Linguistics*, 15, 351–396.

O. E. D. (1989). *Oxford English Dictionary*. Electronic edition (also available online).

Petré, P. (2012). General Productivity: How become waxed and wax became a copula. *Cognitive Linguistics*, 23, 28–65. https://doi.org/10.1515/cog-2012-0002

Petré, P. (2014). *EEBO Corp Concodancer 1.5*. https://lirias.kuleven.be/handle/1234 56789/436853.

Plag, I. (1999). *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110802863

Schröder, A. (2011). *On the productivity of verbal prefixation in English: synchronic and diachronic perspectives*. Tübingen: Narr Verlag.

Serra Borneto, C. (1996). *Liegen* and *stehen* in German: A study in horizontality and verticality. In E. Casad (Ed.), *Cognitive Linguistics in the Redwoods* (pp. 458–505). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110811421.459

Talmy, L. (2000). *Toward a Cognitive Semantics*. Volume 1. Cambridge: MIT Press.

Torres Cacoullos, R., & Walker, J. A. (2009). The present of the English future: Grammatical variation and collocations in discourse. *Language*, 85, 321–354. https://doi.org/10.1353/lan.0.0110

Traugott, E. C. (2014). Toward a constructional framework for research on language change. *Cognitive Linguistic Studies*, 1, 3–21. https://doi.org/10.1075/cogls.1.1.01tra

Van Oosten, J. (1984). Sitting, Standing, Lying in Dutch: A Cognitive Approach to the Distribution of the Verbs Zitten, Staan, and Liggen. In J. Van Oosten, & J. Snapper (Eds.), *Dutch linguistics at Berkeley* (pp. 137–160). Berkeley: UCB.

Yamagushi, T., & Newman, J. (2002). Action and state interpretation of 'sit' in Japenese and English. In J. Newman (Ed.), *The Linguistics of Sitting, Standing and Lying* (pp. 43–59). Amsterdam: John Benjamins.

Pseudocoordination in Norwegian

Degrees of grammaticalization and constructional variants

Torodd Kinn
University of Bergen

Pseudocoordination in the Scandinavian languages is often claimed to involve grammaticalization. Such coordination is exemplified by Norwegian *sitte og lese* 'be reading' (literally: 'sit and read'), *gå og se* 'go and see', and *løpe rundt og leke* 'run around and play', where the first verb in some cases expresses partly or even mainly grammatical meaning. Using a large contemporary corpus, the article explores constructional variants involving atelic motion, like *løpe rundt* 'run around'. Indicators of grammaticalization such as paradigmatization, bleaching, and decategorialization of the first verb are in focus, as well as the concomitants of grammaticalization in the larger construction. Some first verbs do show signs of grammaticalization, but this is not a general characteristic of the construction.

Keywords: Norwegian, pseudocoordination, verb, grammaticalization, paradigmatization, bleaching, decategorialization, schema, subschema

1. Introduction[1]

This study has an empirical and a theoretical purpose. Empirically, it explores constructional variants of Norwegian pseudocoordination (a kind of asymmetric verb phrase coordination; for alternative terms found in the literature, see Wiklund, 2007, p. 9) and the degrees of first-verb grammaticalization that they exhibit. The constructional variants in focus all denote atelic motion. In theory and methodology, the study combines ideas and methods from grammaticalization theory

1. I thank my employer, the Department of Linguistic, Literary and Aesthetic Studies at the University of Bergen, for funding the research on which this paper is based. I further wish to thank the editors and several reviewers for their advice, which has helped me improve this text considerably.

https://doi.org/10.1075/cal.21.c4

(see Section 3) and constructional approaches to grammar (see Section 2). The relation between grammaticalization and the more general notion of constructionalization is a central theme in such recent works as Traugott & Trousdale (2013) and Trousdale (2014). Grammaticalization is a diachronic phenomenon whose synchronic reflexes can be detected in the degrees of grammaticalization of the constituents of constructions at different levels of generalization: from the most abstract schemas via subschemas to microconstructions. The present study builds on synchronic corpus data and has its focus on degrees of grammaticalization in present-day Norwegian – an approach made necessary by the lack of sufficient diachronic data. Textually, the constructional variants and their properties are in the foreground throughout, while the combination of theoretical approaches is more often demonstrated than discussed.

Pseudocoordination resembles ordinary verb phrase coordination, but there is an asymmetric relation between the verb phrases. The distinguishing characteristics of the construction are presented in Section 2, but two examples are given in (1) and (2):

(1) Jeg *satt* og *så* nedover på den grønne dalen (LBK)[2]
I sit.PST and see.PST downwards on the green the.valley
'I was (/sat) looking down on the green valley'

(2) Raskt *løp* han ned til bekken og *fylte* flaska med vann (LBK)
quickly run.PST he down to the.creek and fill.PST the.bottle with water
'Quickly, he ran down to the creek and filled the bottle with water'

Most kinds of pseudocoordination are common to the Mainland Scandinavian languages (Danish, Norwegian, and Swedish), and there has been a good deal of research on them (e.g. Vannebo, 1969, 2003; Faarlund et al., 1997, pp. 534–535; Tonne, 1999, 2001; Lødrup, 2002, 2014; Hesse, 2009; Bjerre & Bjerre, 2007; Kjeldahl, 2010; Nielsen, 2011; Hansen & Heltoft, 2011, Chapter VIII; Josefsson, 1991; Teleman et al. 1999, volume 4, pp. 902–909; Wiklund, 2007; Kvist Darnell, 2008; Hilpert & Koops, 2008; Ahlberg et al., 2015; Blensenius, 2015; Kuteva, 1999, 2001; see Kvist Darnell, 2008, for more pre-2008 references). Similar constructions are found also in some other Germanic languages (for Faroese, see Heycock & Petersen, 2012; for Low German, see Höder, 2012; for Afrikaans, see de Vos, 2005; for English, see e.g. Stefanowitsch, 2000; Hopper, 2002; Newman & Rice,

2. Most examples are from the *Leksikografisk bokmålskorpus* (LBK), see Section 4. Verbs used in pseudocoordination are in italics in the Norwegian examples and are morphologically glossed. Other words are given a word-based glossing. Glossing abbreviations: IMP = imperative, INF = infinitive, PTCP = perfect participle, PRS = present tense, PST = past tense.

2004, 2008; de Vos, 2005; Wulff, 2006; Chaves, 2007; Bachmann, 2013; Lesuisse & Lemmens, this volume).

Two themes tend to recur in studies on pseudocoordination: First, is it a type of coordination, subordination, or something else? (See e.g. Johnsen, 1988; Lødrup, 2002; Wiklund, 2007.) Second, are the first verbs ordinary verbs or of some (semi)grammaticalized verb type like auxiliaries or light verbs? (See e.g. Kuteva, 1999, 2001; Hesse, 2009.)

In the literature on Scandinavian languages, the term *pseudocoordination* is commonly used, and I have chosen to keep it even though with the prefix *pseudo-* it is arguably a misnomer. The term is based on the view that this is not proper coordination, mainly because the Coordinate Structure Constraint (CSC; Ross, 1967) may be violated (see Section 2.3). However, Lakoff (1986) argues convincingly for English that the possibility of such violation is semantically determined, and examples of pseudocoordination are regarded as "real" coordination (see also Chaves, 2007). I believe this to be tenable for Norwegian (Scandinavian) pseudocoordination, too; it typically denotes naturally cooccurring events. See further Section 2.4.

The question of grammaticalization (auxiliation) relates to the degree to which the first verbs are bleached of lexical meaning (desemanticized) and have come to express (primarily) aspect/aktionsart or some similar grammatical concept. (Note that the first verb in (1) above is most naturally left out in the English translation.) This question is typically tied to the matter of paradigmatization of the first-verb slot, since it has repeatedly been claimed that only a small number of verbs may be used as first verbs (see Section 2.2). Also, the first verb may be somewhat decategorialized (see Section 3.4). (See Hopper (1991) and Lehmann (2015) on the mentioned indicators of grammaticalization: bleaching, paradigmatization, and decategorialization.)

The present study examines the degree of grammaticalization for first verbs in a subschema of pseudocoordination (viz. atelic motion) using collostructional analysis, a corpus-based method from constructional approaches to grammar (e.g. Stefanowitsch & Gries, 2003). While finding that at least one verb (*gå* 'go, walk') exhibits clear signs of bleaching and decategorialization, the study shows that a large number of verbs may be first verbs in this subschema, and that the notion of a small first-verb paradigm for the pseudocoordination schema is misleading.

2. Pseudocoordination as a construction

Using Traugott & Trousdale's (2013, pp. 11–13) three dimensions for the description of constructions, pseudocoordination may be characterized as (a) rather highly complex, involving two potentially complex VPs and their external relations, (b) almost fully schematic, with the coordinator *og* 'and' as the only substantive expression, and (c) mostly contentful (lexical), involving verbs and their complements and modifiers, but with procedural (grammatical) elements, notably the coordinator, precedence relations, and agreement.

2.1 Overall structure

The pseudocoordination schema is itself a subschema of verb phrase coordination. It basically consists of two verb phrases (VP1, VP2) coordinated by *og* 'and'. The verbs are in the same inflectional form.[3] Each verb phrase is normally verb-initial, with complements and modifiers following the verb. (Infrequently, there may be an adverbial constituent in front of the verb.) The postverbal slot complexes will be referred to as postfields. Thus, pseudocoordination consists of a first verb (V1), an optional postfield (PF1), the coordinator (Co), a second verb (V2), and an optional postfield (PF2). This is illustrated in (3)–(4) with nonfinite pseudocoordinations and (5)–(6) with finite pseudocoordinations in embedded clauses. While (3) and (5) have constituents in both postfields, (4) and (6) have none. (Subscripts and brackets are provided only in this section. Note that the brackets do not necessarily enclose single constituents.)

(3) Jeg liker ikke å *gå* $_{V1}$ [omkring]$_{PF1}$ og$_{Co}$ *stinke* $_{V2}$ [røyk heile tida]$_{PF2}$
I like not to go.INF around and reek.INF smoke the.whole the.time
'I don't like to go around reeking of smoke all the time'

(4) Nå må du *komme* $_{V1}$ og$_{Co}$ *se* $_{V2}$
now must you come.INF and see.INF
'Now you must come and have a look'

3. Finite forms in Norwegian are the present tense, the past tense, and the imperative. Nonfinite forms are primarily the infinitive and the perfect participle. (In pseudocoordination, the perfect participle is in the supine form (singular, neuter, nondefinite), cf. Faarlund et al. (1997, p. 473).) A present participle V1 is in Norwegian regularly followed by an infinitive V2, see Kinn (2014, 2017), i.e. the verbs do not agree. See Blensenius (2015) for exceptions to verb agreement in Swedish.

(5) Jeg er glad for at jeg ikke *går* $_{V1}$ [omkring]$_{PF1}$ og$_{Co}$ *stinker* $_{V2}$ [røyk heile tida]$_{PF2}$
I am glad for that I not go.PRS around and reek.PRS smoke the.whole the.time
'I'm glad that I don't go around reeking of smoke all the time'

(6) Jeg krever at du *kommer* $_{V1}$ og$_{Co}$ *ser* $_{V2}$
I demand that you come.PRS and see.PRS
'I demand that you come and have a look'

Finite pseudocoordinations in matrix clauses may be discontinuous. Matrix clauses display verb-second syntax, which means that V1 follows one initial topicalized constituent (except in yes–no questions and most imperatives, where no constituent is topicalized). Constituents that are neither topicalized nor belong in PF1 will then appear between V1 and PF1. Common intervening constituents are subjects and so-called sentence adverbials (mostly epistemic and evaluative disjuncts and connectives, but also e.g. *ikke* 'not'). These are semantically external to the pseudocoordination, but the subjects are shared by V1 and V2, and sentence adverbials have scope over the whole coordination. The slot complex between V1 and PF1 will be referred to as the midfield (MF). A matrix clause with the midfield and both postfields filled is exemplified in (7), while (8) has no constituent in any of these fields.

(7) Heldigvis *går* $_{V1}$ [jeg ikke]$_{MF}$ [omkring]$_{PF1}$ og$_{Co}$ *stinker* $_{V2}$ [røyk heile tida]$_{PF2}$
luckily go.PRS I not around and reek.PRS smoke the.whole the.time
'Luckily, I don't go around reeking of smoke all the time'

(8) *Kom* $_{V1}$ og$_{Co}$ *se* $_{V2}$!
come.IMP and see.IMP
'Come and have a look!'

2.2 The verb slots

V2 may apparently be almost any verb, which is followed as appropriate by objects, predicate complements and adverbials in PF2. Even though there appear to be few restrictions on the choice of V2, the use of V2s provides crucial information about the status of V1 (see Section 3.5).

The V1 slot is the point where it has been assumed that pseudocoordination is paradigmatically restricted; it is often claimed that the number of possible V1s is small (see Section 3.3). Some researchers (e.g. Teleman et al., 1999, volume 4,

pp. 902–909) have proposed taxonomies of pseudocoordination types based on VP1 semantics. In the terms of Traugott & Trousdale (2013), if pseudocoordination is treated as a schema, such taxonomies describe subschemas at different levels of abstraction. Clusters of semantically related V1s then define the nodes of the various taxa.

With slight modifications of the taxonomy of Teleman et al. (1999) for Swedish, the central types of pseudocoordination may be listed as follows, with Norwegian examples of V1:

1. location
   a. posture (*ligge* 'lie', *sitte* 'sit', *stå* 'stand')
   b. being elsewhere (*være* 'be')
   c. atelic motion (*gå (rundt)* 'go (around)', *løpe omkring* 'run around')
2. telic motion (*gå (ut, bort, …)* 'go (out, away, …)', *løpe (ned, av sted, …)* 'run (down, off, …)')
3. phase (*holde på* 'carry on')
4. communicative channel (*ringe* 'phone')
5. kindness (*være snill* 'be (so) kind')

Note that terms such as *telic* and *atelic* here refer to the meaning of VP1, not necessarily the whole pseudocoordination.

Although Teleman et al. (1999) is among the works that recognize a wider range of pseudocoordination than has been common, it is becoming clear that the actual range is wider. Lødrup (2002) mentions V1s of assuming a posture (*sette seg* 'sit (down)'), and I can add pseudocoordination where V1 is a verb of halted motion (*stanse* 'stop'), body-part motion (*rote rundt* 'poke around'), visual inspection (*se seg rundt* 'look around'), and other states than kindness (*være heldig* 'be lucky'). The list will presumably be extended as research on pseudocoordination progresses, see e.g. Ahlberg et al. (2015).

Most V1s are intransitive (although some do take a reflexive object, e.g. *sette seg* 'sit (down)'), which means that most constituents in PF1 are adverbials, most typically of place or direction.[4] Copular V1 complexes like *være snill* 'be (so) kind' have a predicate complement (e.g. *snill*) in PF1, and free secondary predicates may also be found in PF1. Also, for many V1s, pseudocoordination can be combined with the presentational (existential) construction, in which case there is a presented subject in PF1.

4. There are at least some idiomatic transitive V1 complexes, like *ta mot til seg* 'pick up the courage' (literally: 'take courage to oneself') and *ta bladet fra munnen* 'speak out' (literally: 'take the leaf from the mouth'). The field of transitive V1s is largely unexplored.

2.3 Pseudocoordination and ordinary VP coordination

Much of what has been said above about pseudocoordination also holds for ordinary verb phrase coordination. Good overviews of the criteria mentioned in the literature for distinguishing pseudocoordination from ordinary VP coordination are given in Teleman et al. (1999) and Hesse (2009, pp. 13–16).[5]

In ordinary VP coordination there are no restrictions on the choice of V1, and the coordinated verbs do not have to agree (although they typically do). The restrictions on V1 and lack of restrictions on V2 in pseudocoordination are part of an asymmetry; the VPs are in a fixed sequence and cannot be interchanged (cf. Newman & Rice, 2008, for asymmetry in English). There is no such restriction on ordinary VP coordination. Further, *og* 'and' is the only coordinator that can be used in pseudocoordination; differently from ordinary VP coordination, neither *eller* 'or' nor *men* 'but' are used, nor can *og* 'and' be preceded by *både* 'both' or *henholdsvis* 'respectively'.

Another difference concerns sentence adverbials. In pseudocoordination, a sentence adverbial placed in relation to V1 also has scope over V2, differently from ordinary coordination. For instance, *dessverre* 'unfortunately' in (9) evaluates the whole situation of the subject referent's sitting inside and keeping to herself, while in the ordinary coordination in (10) it is the subject referent's sulking that is evaluated and not the isolation.

(9) Hun *sitter* dessverre inne og *isolerer* seg
she sit.PRS unfortunately inside and isolate.PRS her
'Unfortunately, she's (sitting) inside keeping to herself'

(10) Hun furter dessverre over noe og isolerer seg
she sulks unfortunately over something and isolates her
'She's sulking about something, unfortunately, and keeping to herself'

A clear difference between pseudocoordination and ordinary coordination is that only the former allows violation of the CSC (Ross, 1967): A dependent of V2 that is not also a dependent of V1 may (for instance) be topicalized, i.e. it can, in transformational terms, be moved out of the coordination. Ordinary coordination allows topicalization of a constituent that is a dependent of both V1 and V2, as shown in (11), but not CSC violation as in (12). (13) illustrates pseudocoordination with CSC violation.

5. Some of the criteria are actually better suited to distinguish verb coordination from VP coordination, including pseudocoordination. The details are not of further concern here.

(11) Han har jeg møtt og snakka med
him have I met and talked with
'Him, I've met and talked to'

(12) *Han har jeg møtt begge to og snakka med
him have I met both two and talked with
(*Him, I've met both of them and talked to')

(13) Han har jeg *sittet* lenge og *snakka* med
him have I sit.PTCP long and talk.PTCP with
'Him, I've been (sitting) talking to for a long time'

CSC violation is not limited to topicalization, but can be found also in matrix wh-clauses (*Hvem satt du og snakka med?* 'Who were you (sitting) talking to?'), embedded wh-clauses (*Jeg spurte hvem du satt og snakka med* 'I asked who you were (sitting) talking to'), clefting (*Det var han jeg satt og snakka med* 'It was him that I was (sitting) talking to'), and relativization (*den mannen som jeg satt og snakka med* 'the man that I was (sitting) talking to').

2.4 Constructional meaning

It has been noted at least since Huldén (1961) that pseudocoordination refers to specific events rather than general properties. Thus, while the simple *Hun skreiv brev* 'She wrote letters' may (also) be interpreted as referring to a habitual activity, the pseudocoordination *Hun satt og skreiv brev* 'She was (sitting) writing letters' must be understood as referring to a particular writing occasion.

Another commonly observed property of pseudocoordination is that it has a single-event meaning, e.g. Josefsson (1991), Lødrup (2002), and Wiklund (2007). In my view, this should not be understood to the effect that the meaning of the whole is the same as the meaning of VP2; rather it is inherent to the meaning of the whole that two subevents – the meanings of VP1 and VP2 – are construed as forming a tightly integrated complex event. Other distinguishing properties can also be understood in this light. Only the additive coordinator *og* 'and' can be used (see Section 2.3); other coordinators would indicate that the subevents should be understood as more clearly separated. Further, sentence adverbials have scope over both VP1 and VP2, treating them as one.

It might be objected against single-event meaning that pseudocoordination with V1s of telic motion (e.g. *gå ut og kjøpe* 'go out and buy') refer to sequential events (see Newman & Rice, 2008, for English). They do, but the construction construes them as forming one complex event. One consequence of this construal

is that there can be no temporal adverbial in VP2 that locates it after VP1 in time. Compare (14) and (15):

(14) Hun *gikk* i butikken og *kjøpte* poteter
she go.PST in the.shop and buy.PST potatoes
'She went to the shop and bought potatoes'

(15) Hun gikk i butikken og kjøpte så poteter
she went in the.shop and bought then potatoes
'She went to the shop and then bought potatoes'

(14) is a pseudocoordination; the subject referent's going to the shop is intimately tied to her buying potatoes; she bought the potatoes in the shop. (See Nielsen (2011) for discussion of shared locatives.) In (15), which represents ordinary coordination, the adverb *så* 'then' ('thereafter') indicates that the buying of potatoes constitutes a separate event, presumably not located in the shop.

The single-event meaning of the whole pseudocoordination is closely connected to the meaning relation between VP1 and VP2. For English, Lakoff (1986) describes an example like *go and buy* as a natural course of events. I would generalize this: Pseudocoordination denotes naturally cooccurring events, sequential or simultaneous. A somewhat different characterization of pseudocoordination is given for Danish by Nielsen (2011), who uses the term *muliggørelse* 'making possible' about the VP1–VP2 relation. In (14) above, the subject referent's going to the shop makes it possible for her to buy potatoes there. I believe Nielsen's interpretation is on the right track, but 'making possible' is too strong; 'making more feasible' is more adequate. For instance, in *sitte og lese* (literally: 'sit and read') the sitting posture is not a necessity for the reading, but it is convenient. Thus, *facilitation* might be a better term for the relation between VP1 and VP2. For instance, in *gå rundt og samle inn penger* 'go around collecting money' the motion facilitates the collection, in *ringe og bestille pizza* 'phone and order pizza' the use of the phone facilitates the ordering, etc.

3. Pseudocoordination and grammaticalization

Pseudocoordination has often been regarded as involving some degree of grammaticalization. Kvist Darnell (2008, p. 44) finds that a majority of modern descriptions agree that the main function of pseudocoordination is to provide, strengthen or maintain aspect/aktionsart. Hesse (2009) has as its main title *Zur Grammatikalisierung der Pseudokoordination*. Thorell (1973) describes V1 in some pseudocoordination types as a kind of auxiliary with a weakened meaning,

categorizing them into those with a durative/imperfective meaning (mainly posture verbs) and those which signify the sudden beginning of an event (mainly verbs of telic motion). Faarlund et al. (1997, pp. 647–648) include pseudocoordination with posture verbs under constructions with cursive aspect. Tonne's (2001) dissertation on progressives in Norwegian has pseudocoordination with posture verbs and *drive* 'carry on' as central objects of study, and posture verb pseudocoordination is in focus in Kuteva's (1999, 2001) work on auxiliation. The exact aspect/aktionsart contribution of (some subschemas of) pseudocoordination is not in focus here; for this I refer to Tonne (2001) and Hesse (2009), and see further Blensenius (2015) for a critical appraisal of the view that posture pseudocoordinations in Swedish are progressives.

The most strongly grammaticalized part of pseudocoordination is arguably the coordinator *og* 'and', but it is members of the V1 slot that have been discussed as grammaticalized in the literature. As emphasized by Himmelmann (2004), grammaticalizing elements do not grammaticalize independently of context. Other parts of the construction are also affected. For pseudocoordination, I will discuss the following below: (a) the coordinator, (b) the meaning or function of V1s, (c) the paradigm of V1s, (d) the constituents in the postfields, (e) the semantic relation between V1 and V2, and (f) the meaning of the construction as a whole.

The present study approaches these issues on the basis of contemporary synchronic language data. Although there have been observations about the degree of grammaticalization in Scandinavian pseudocoordination, the actual diachronic development has rarely been addressed. Exceptions are Vannebo (2003) on Norwegian *ta* 'take' and Hilpert & Koops (2008) on Swedish *sitta* 'sit' (see Section 3.4 for the latter). While diachronic data are better sources for knowledge of past developments, they may not always be available to an extent or in a format that would have made studies possible (see e.g. Hilpert, 2010). The present study relies on the availability of a large corpus that has been lemmatized and tagged for parts of speech and inflection. No such corpora exist for the older stages of Norwegian.

However, the gradual nature of grammaticalization often makes diachrony observable in synchrony, since older variants tend to live on alongside the younger variants that have diverged from them (Hopper 1991). This is central in Heine et al's. (1991, pp. 258–261) reflections on panchrony and the impracticality of too strict a separation of synchrony and diachrony. The comparison of constructions with different synchronic degrees of grammaticalization does to some extent allow hypotheses about past developments.

In the case of pseudocoordination, the elements assumed to be affected by grammaticalization are synchronically present in several variants. It is in most cases the comparison of synchronic variants that has allowed researchers to observe the differences and conclude that V1s in some subschemas have been

grammaticalized to a higher degree than others. Some V1s are seen to have a less lexical meaning in pseudocoordination than elsewhere, the postfields of some V1s are seen to be less frequently filled in pseudocoordination than when the verbs are used in other constructions; etc. Still, the limitations of synchronic data need to be acknowledged, and I therefore refer to synchronic degrees of grammaticalization rather than make claims about diachrony.

3.1 The coordinator

The set of coordinators is in itself a small paradigm, but the paradigm in pseudocoordination has only one member: *og* 'and'. This is related to the single-event meaning of the construction (see Section 2.4). In some conceptions of grammaticalization (e.g. Givón, 1979), the loss of the grammaticalized element is seen as the extreme endpoint, and indeed, *og* may be left out in at least some types of pseudocoordination in certain Norwegian dialects (Faarlund et al., 1997, p. 535). Pseudocoordination without the coordinator formally satisfies Aikhenvald's (2006) definition of serial verb constructions. In fact, writers such as Nielsen (2011) and Hansen & Heltoft (2011) regard pseudocoordination with *og* as serialization. The issue will not be elaborated here, but pseudocoordination and serialization may be regarded as closely related syndetic and asyndetic construction types.

When V1 is strongly bleached, the additive meaning of *og* becomes blurred, since VP1 contributes little to the propositional content of the whole (see Section 3.6).

3.2 V1 meanings and functions

It is above all posture V1s that have been discussed in terms of grammaticalization in pseudocoordination (e.g. Kvist Darnell, 2008), and posture verbs are a well-known grammaticalization source (Heine & Kuteva, 2002). Posture verbs have been studied extensively also in other constructions, like Holm (2013), Lemmens (2005), Newman & Rice (2004), e.g. with respect to what kinds of nonhuman objects tend to be regarded as lying, sitting and standing, respectively.

Attention to posture verbs in pseudocoordination predates modern grammaticalization theory. Vannebo's (1969) central work on aktionsart in Norwegian points to a certain degree of bleaching in some subschemas. Others, however, have rightly pointed out that when for instance the posture verbs *ligge* 'lie', *sitte* 'sit', and *stå* 'stand' are used, the intended meaning is normally that the subject referent is indeed lying, sitting, or standing, respectively, while performing the action denoted by V2 (e.g. Tonne, 2001; Lødrup, 2002). A key to understanding this disagreement is probably the concept of 'backgrounding' (Chaves, 2007). When the

verb *sitte* 'sit' is used, for example, it is referentially appropriate in the sense that the subject referent is sitting, but it is informationally less important than V2. This ties in with the notion of a single-event meaning; the state of affairs referred to with V1 is treated as a backgrounded facilitator for the state of affairs referred to with V2. Consider (16).

(16) De sier at jeg bare *sitter* og *ser* på TV, men det er ikke sant. Jeg leser
they say that I just sit.PRS and see.PRS on TV but that is not true I read
mye.
much
'They say that I just sit and watch TV, but that's not true. I read a lot.'

Denying the truth of 'just sit and watch TV' normally amounts to denying 'just watch TV'. (Denying 'sit' is possible, but will usually have a comical effect.) Thus, although posture V1s are typically not clearly bleached, their meaning is backgrounded and less addressable (cf. Boye & Harder, 2012) than that of V2s.

Somewhat paradoxically, maybe, the V1s that clearly do exhibit bleaching have received less attention. For Swedish, Kvist Darnell (2008, p. 106) identifies these as grammaticalized V1s, and her list comprises *ta* 'take', *gå* 'go, walk' (both atelic and telic motion, see Section 2.2), and *komma* 'come'. For Norwegian I would add at least *drive* 'carry on', cf. Tonne (2001), Lødrup (2002). Examples where V2 shows that V1 cannot be understood quite literally are clear evidence of bleaching. As an example, atelic *gå* 'go, walk' can be used as in (17), where it clearly cannot be the case that the subject referent had continually been going about on her feet for ten years.

(17) En mor hadde i 10 år *gått* og *sørget* over sønnen sin (LBK)
a mother had in ten years go.PTCP and grieve.PTCP over the.son her
'A mother had been (/gone around) grieving her son for ten years'

Hesse (2009, pp. 92–99) discusses the aspectual properties of such atelic pseudocoordination with *gå*. An example with *drive* 'carry on' is given in (18), where V1 has very little independent meaning but appears to underscore the active nature of the event and add a slightly pejorative evaluation of the activity referred to with V2.

(18) Et engelsk tv-team *driver* og *intervjuer* en dresskledd palestiner …
an English TV.team carry.on.PRS and interview.PRS a suit.dressed Palestinian
'An English TV team is interviewing a Palestinian in a suit' (LBK)

This use of *drive* is found only in pseudocoordination, and without *og* and VP2 the expression would be ungrammatical. Apparently, *drive* is here a reduction of the still extant particle verb *drive på* (*på* means 'on'), as in (19):

(19) Så lenge supermaktene *drev* på og *sloss*, trodde [vi] … (LBK)
as long the.super.powers carry.on.PST on and fight.PST believed we
'As long as the superpowers were (/kept) fighting, we believed …'

The use of *drive på* is possible also outside of pseudocoordination (see Section 3.4), which makes it probable that it is the source of the use without *på*.

3.3 V1 paradigmatization

It has been claimed about pseudocoordination at the schematic level that the number of V1s is small, e.g. "a limited number of verbs of (mostly) location or direction" (Johannessen, 1998, pp. 48–49), "a small number of first verbs" (Lødrup, 2002, p. 121), a "[r]estricted set of verbs" (Wiklund, 2007, p. 98). However, some works, like Teleman et al. (1999) and Hansen & Heltoft (2011), have listed or exemplified more than 30 different V1s. As shown by the taxonomy from Teleman et al. (1999) and the existence of additional types referred to in Section 2.2, several subschemas may be recognized. Some of these may exhibit rather small paradigms of V1s; e.g., posture verb pseudocoordination appears to be limited to four verbs: *ligge* 'lie', *sitte* 'sit', *stå* 'stand', and the rather infrequent *henge* 'hang'.[6] The set of V1s of halted motion is probably also small, but as yet unexplored. On the other hand, V1s of atelic motion will be shown in Section 4 to be numerous.

Not only the number of members of a V1 (subschema) paradigm is relevant, however. Some members may be more central than others (see Bybee, 2010, especially pp. 88–90), in terms of e.g. frequency and of signs of grammaticalization. In other words, the paradigms are structured. It will be shown that although *gå* 'go, walk' is only one of many possible atelic motion V1s, it is undeniably the central member in terms of frequency and one of only few atelic motion verbs to exhibit bleaching.

3.4 Constituents in the postfields

Bleaching of V1 is connected to a tendency for PF1 to host fewer constituents, observed by Hesse (2009). Hilpert & Koops (2008) find in their diachronic study of Swedish pseudocoordination with *sitta* 'sit' that PF1 has over time come less frequently to be filled with adverbial constituents. They regard this as evidence of decategorialization of the verb in this use, since *sitta* outside of pseudocoordination is usually combined with a locative adverbial. That is, the argument structure

6. This is probably related to the fact that the basic posture verbs in themselves form a small paradigm, i.e., paradigmaticity is not specific to pseudocoordination.

of the verb is changing. Also, Hilpert & Koops (2008) find an increasing tendency for locative adverbials belonging most naturally in PF1 to be placed in PF2. A Norwegian example is given in (20):

(20) Jeg *satt* og *leste* i en behagelig stol … (LBK)
I sit.PST and read.PST in a comfortable chair
'I was (/sat) reading in a comfortable chair'

(The alternative *Jeg satt i en behagelig stol og leste* is also well-formed.) While *jeg satt i en behagelig stol* 'I was sitting in a comfortable chair' sounds quite natural, *jeg leste i en behagelig stol* 'I was reading in a comfortable chair' is grammatical, but sounds awkward – i.e., the locative is semantically more closely related to V1 *sitte* 'sit' than to V2 *lese* 'read'. The diachronic tendency found by Hilpert & Koops (2008) may be interpreted in the light of the single-event meaning of all pseudocoordination: PF2 is treated as common to V1 and V2.

This tendency is taken further by the bleached *drive* 'carry on'; it is ungrammatical to place a locative adverbial in PF1 rather than PF2.[7] (21) gives an example:

(21) at det ofte er politi som *driver* og *patruljerer* utafor klubben (LBK)
that it often is police that carry.on.PRS and patrol.PRS outside the.club
'that there are often police patrolling outside the club'

The locative *utafor klubben* could not have been placed between *driver* and *og*. In addition to having lost *på* (see Section 3.2), *drive* has lost its ability to relate directly to location. The assumed source *drive på* has no such limitation, as shown in (22):

(22) Om ettermiddagen dreiv han på i smia (LBK)
about the.afternoon carried.on he on in the.smithy
'In the afternoon, he was carrying on (/working) in the smithy'

3.5 The semantic relation between V1 and V2

As described in Section 2.4, the meaning relation between V1 and V2 is one of facilitation: The posture, motion etc. of V1 prepares the ground, so to speak, for the activity or action of V2. But when V1 is bleached, the relation of facilitation becomes obscured. Examples (17)–(18) above demonstrate this; in *gått og sørget* 'been grieving' (literally: 'gone and grieved') and *driver og intervjuer* 'is interviewing' (literally: 'carries on and interviews') the V1s contribute little propositional

7. It is possible (though not common) to have certain other kinds of dependent constituents after *drive* in pseudocoordination, e.g. manner adverbials and secondary predicates.

meaning. The remaining meaning may be primarily aspectual (cf. Hesse (2009), Kvist Darnell's (2008, p. 44) summary of previous research, and Blensenius's (2015) critical discussion).

3.6 The meaning of the construction as a whole

The meaning of an ordinary coordination is that of a whole composed of (at least) two parts (see Wierzbicka, 1980, Chapter 7). In additive coordinaton with 'and', the parts are construed as participating in the same way with the elements outside of the coordination. As described by Langacker (1991, pp. 472–490), it varies how salient the whole and the parts are relative to each other.

The observation in Section 2.4 that pseudocoordination construes the meanings of two coordinated VPs as referring to one complex event entails that the parts are seen as closely related (by facilitation), and this strengthens the salience of the whole. When V1 is bleached and even decategorialized, one of the parts of the coordination becomes even less salient, and the whole is little more than the part referred to with VP2 in terms of propositional content.

3.7 Summary

As regards the grammaticalization of pseudocoordination, the following points have been observed above:

- The coordinator *og* 'and' is highly grammatical and invariable (and may sometimes be omitted in some dialects). V1 bleaching obscures the additive meaning of *og*.
- Some V1s are bleached and have mostly grammatical (procedural) meaning, but this is not a general property of the pseudocoordination schema.
- At the schema level, there is little V1 paradigmatization, but certain subschemas have small paradigms.
- Scarcity or absence of constituents in PF1 appears to be an indicator of V1 grammaticalization. The same goes for the placement of V1 dependents in PF2.
- Bleaching of V1 causes the relation of facilitation between V1 and V2 to be blurred.
- Bleaching of V1 causes the meaning of the coordination to approach the meaning of V2, thus strengthening the single-event meaning of the construction.

4. Pseudocoordination of atelic motion

This part of the article is based on two closely related corpus studies of contemporary Norwegian Bokmål. The corpus employed is *Leksikografisk bokmålskorpus*, a balanced corpus of written Bokmål from the period 1985–2013 containing about 100,000,000 words. The corpus is lemmatized and tagged for parts of speech and inflectional morphology.[8] The focus of the studies is on one of the subschemas of the taxonomy delineated in Section 2.2, viz. pseudocoordination of atelic motion.

The first study targets a variant where there is an adverb in PF1 indicating the atelicity of motion, viz. *rundt* or *omkring*, both 'around, from place to place', e.g. *gå rundt og nynne* 'go around humming'. (This is parallel to one of the variants explored for English by Stefanowitsch (2000).) A subdivision is made between variants with and without additional words[9] between the atelicizing adverb and *og*.

The second study targets another variant of atelic motion, namely one where there is no atelicizing adverb, e.g. *gå og nynne* 'go (around) humming'. It is known (e.g. Josefsson, 1991) that the variant with an atelic adverb allows more different V1s than the one without, but there has to date been no empirical study of this. In the following text, the abbreviations shown after the constructional variants illustrated in (23)–(25) will be used. In addition, I will use *Vr(X)oV* as a cover term for *VrXoV* and *VroV*.[10]

(23) *gå* rundt i huset og *rydde* (VrXoV)
go.INF around in the.house and tidy.INF
'go around the house tidying up'

(24) *gå* rundt og *rydde* (VroV)
go.INF around and tidy.INF
'go around tidying up'

(25) *gå* og *rydde* (VoV)
go.INF and tidy.INF
'go (around) tidying up'

A central aim of the first study is to document the range of V1s that are actually used in Vr(X)oV, i.e. the size of the V1 paradigm of this subschema. The study fur-

8. See http://www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/skriftsprakskorpus/lbk/index.html.

9. These may or may not form a constituent together with *rundt/omkring*.

10. *r* is mnemonic for *rundt/omkring*, *o* is for *og*, and *X* is for words between *rundt/omkring* and *og*.

ther compares the most frequent V1s with respect to two indicators of degree of grammaticalization: (a) the proportions of VroV and VrXoV – lower frequencies of X constituents are assumed to indicate higher degrees of grammaticalization–, and (b) the V2 collexemes of the V1s – certain V2 types (notably of cognition) may indicate V1 bleaching, another aspect of grammaticalization. A collexeme is a lexeme that tends to collocate with a certain construction or other lexeme (Stefanowitsch & Gries, 2003). Collexemes are identified through the use of collostructional analysis, see below.

The second study is more limited and targets VoV structures, i.e. those where PF1 is empty. Assuming that less material between *rundt/omkring* and *og* in Vr(X)oV indicates a higher degree of grammaticalization, VoV is an even denser formal type and can be expected to exhibit higher degrees of grammaticalization in another respect, viz. V1 bleaching as indicated by V2 collexemes. The study takes as its point of departure the more frequent V1s of Vr(X)oV: (a) To what extent are the V1s used also in VoV, and (b) do their V2 collexemes provide evidence of stronger V1 bleaching in VoV than in Vr(X)oV?

Collostructional analysis (see Stefanowitsch, 2013 for an overview) was developed as a set of quantitative corpus-linguistic methods in the context of synchronic construction grammar, but it has been extended to diachronic investigation, including grammaticalization studies. For an early example of such fruitful combination of construction grammar and grammaticalization theory, see Hilpert (2008) on Germanic future constructions. The present contribution takes a similar approach.

4.1 First corpus study: Vr(X)oV

The first object of this study is to explore in a large corpus the range of V1s employed in Vr(X)oV, with an atelicizing adverb in PF1. While there are several ways to express atelic motion, the central adverbs are *rundt* and *omkring*, both 'around, from place to place'. As noted above, pseudocoordination has almost no substance at the level of the schema. In this subschema, two slots are fixed, namely the atelicizing adverb *rundt* or *omkring* in PF1 and the coordinator *og*. The verb slots are open, but the verbs must agree. PF1 may house other constituents (X) after the atelicizing adverb (infrequently before it), and when V1 is in verb-second position of a matrix clause, there may be constituents in the midfield (esp. subject and sentence adverbials, cf. Section 2.1). Only infrequently is there an adverbial constituent between the coordinator and V2. PF2 cannot be specified.

No maximum numbers of words intervening between V1 and *rundt/omkring*, between *rundt/omkring* and *og*, and between *og* and V2 can be specified, and V1 and V2 are lexically unspecified. This high level of schematicity makes it difficult

to search in large corpora without getting intractable amounts of irrelevant examples. In a corpus of 100,000,000 words, fairly strict limitations need to be put on the number of intervening words.

In the searches, the maximum numbers of intervening words were set to 3 between V1 and *rundt/omkring*, 4 between *rundt/omkring* and *og*, and 3 between *og* and V2. One search was performed per inflectional form: infinitive, past participle, present tense, past tense, imperative.[11] The searches yielded 10,443 hits, out of which 2,728 hits were found to be tokens of Vr(X)oV – i.e. 7,715 hits were irrelevant. The search limitations undoubtedly caused the searches to miss a certain number of examples of the subschema. However, there is little reason to believe that this has skewed the results in any detrimental way relative to the questions posed in the study.

All search hits were downloaded to a spreadsheet and categorized manually. To be categorized as Vr(X)oV, the tokens had to satisfy the following criteria:

- The expression involved coordination of two VPs (rather than noncoordination or coordination of e.g. clauses).
- The coordinator was *og* without *både* 'both' or *henholdsvis* 'respectively'.
- The words *rundt* or *omkring* were adverbs meaning 'from place to place' (rather than prepositions, postpositions, or adjectives, and rather than adverbs with a different meaning, e.g. 'around its own axis').
- VP1 and VP2 referred to time spans construed as simultaneous (rather than sequential). This encompasses 'objective' simultaneity of continuous activities (*hun gikk rundt og sang* 'she went around singing'), a continuous activity in combination with an iterated action (*hun gikk rundt og nøys* 'she went around sneezing'), and a semicontinuous activity alternating with an iterated action (*hun gikk rundt og tente stearinlys* 'she went around lighting candles').
- VP1 could be interpreted as referring to motion of the subject referent. This sometimes required taking VP2 into account, if V1 was not inherently a motion verb. VP2 sometimes indicated bleaching of V1, but such examples were included.
- The motion referred to with VP1 basically involved the subject referent as a whole. The requirement for whole-subject motion excluded a number of constructions, most of which should probably be regarded as closely related subschemas of pseudocoordination, although their nature has not been stud-

11. Constructions with a present participle V1 and an infinitive V2 (see note to Section 2.1) were initially also included, but there were only two tokens. For many verbal present participles, there are formally identical adjectival forms, some highly frequent, and the corpus does not distinguish between them. In light of there being only two relevant tokens, I decided to omit present participles altogether, since this was considered to make the statistics more reliable.

ied in detail. The most important types involved VP1s for body part motion (*rote rundt* 'poke around') – which is sometimes difficult to distinguish from whole-subject motion–, distributed posture (*sitte rundt* 'sit around'; cf. Newman & Rice (2004) for English), establishing visual contact (*kikke seg rundt* 'look around'), establishing communicative contact (*ringe rundt* 'phone around'), and sexual activity (*knulle rundt* 'screw around').

In spite of the strict inclusion criteria described above, the search yielded examples of Vr(X)oV for no less than 167 different V1s. Out of these, 31 occurred at least 10 times, while 60 were hapaxes. There were 673 different V2s; 51 occurred at least 10 times, 360 only once. The relatively high type frequency of 167 V1s and the large number of hapaxes indicate a high degree of syntactic productivity; i.e., we are dealing with an open schema (cf. Barðdal, 2008). The most frequent V1s and V2s are given in Table 1, with frequencies and collostructional strengths (CS). Collostructional strength is a measure of attraction between a construction and a lexeme (Stefanowitsch & Gries, 2003).[12]

Some examples are given in (26)–(28). The most frequent V1, *gå* 'go, walk', is used in (26), which also exemplifies CSC violation (the complement of *på* 'on' in PF2 is relativized). (27) shows one of the V1 hapaxes, the loanword *cruise* 'cruise'. The V1 in (28), *brumme* 'growl, grumble', illustrates the use of a verb (also a hapax) that is not in itself a motion verb, but the combination with *omkring* 'around' coerces it into the meaning 'go (around) growling'.

(26) Da fortjente hun … smerten hun *gikk* rundt og *bar* på (LBK)
then deserved she the.pain she go.PST around and carry.PST on
'Then she deserved the pain she was carrying around'

(27) bruke resten av dagen på å *cruise* rundt og *se* på jentene (LBK)
use the.rest of the.day on to cruise.INF around and see.INF on the.girls
'spend the rest of the day cruising around looking at the girls'

12. The calculation of collostructional strength builds on four figures in a two-by-two contingency table: the frequency of the given verb in the given construction, the frequency of the verb in other constructions, the frequency of other verbs in the construction, and the frequency of other verbs in other constructions. These figures are subjected to Fisher's exact test, which yields a probability value. To aid readability, collostructional strength is conventionally presented as the log10-transformation of the probability. Verbs with an observed frequency higher than the expected frequency receive positive values, while those with a frequency lower than expected are set to negative collostructional strength values. Strength values larger than 1.3 (smaller than −1.3) are statistically significant. Collostructional strength was calculated by using Gries's (2014) R script for collostructional analysis.

Table 1. Most frequent V1s and V2s in the Vr(X)oV subschema

V1	N	%	CS	V2	N	%	CS
gå 'go, walk'	1355	49.67	∞	*være* 'be'	139	5.10	−73.02
reise 'travel'	217	7.95	254.70	*se* 'see, look'	136	4.99	23.25
løpe 'run'	146	5.35	201.82	*lete* 'search'	88	3.23	121.73
vandre 'wander'	93	3.41	168.74	*snakke* 'talk'	55	2.02	23.48
kjøre 'drive'	57	2.09	42.16	*si* 'say'	54	1.98	−0.42
fly 'fly, run'	53	1.94	69.63	*gjøre* 'do'	51	1.87	1.57
springe 'run'	48	1.76	72.78	*ta* 'take'	50	1.83	1.52
dra 'go'	46	1.69	25.10	*tenke* 'think, ponder'	48	1.76	11.01
rusle 'potter'	45	1.65	88.68	*tro* 'think, believe'	42	1.54	10.40
streife 'roam'	20	0.73	34.29	*prøve* 'try'	37	1.36	17.39
være 'be'	20	0.73	−175.58	*leke* 'play (games)'	34	1.25	41.00
farte 'travel'	18	0.66	38.36	*rope* 'call, yell'	34	1.25	25.75
krabbe 'crawl'	17	0.62	30.13	*samle* 'collect'	34	1.25	24.51
virre 'flutter'	17	0.62	36.25	*vente* 'wait'	34	1.25	16.87
drive 'drift'	16	0.59	5.41	*holde* 'hold'	28	1.03	3.61
trave 'trot'	16	0.59	33.45	*få* 'receive'	26	0.95	−5.20
danse 'dance'	15	0.55	14.33	*kikke* 'peek'	25	0.92	20.40
bevege (seg) 'move'	14	0.51	8.58	*selge* 'sell'	25	0.92	17.10
hoppe 'jump'	14	0.51	11.08	*ha* 'have'	24	0.88	−69.21
krype 'crawl'	14	0.51	15.40	*vite* 'know'	24	0.88	0.76
svinse 'flit'	14	0.51	36.22	*huske* 'remember'	22	0.81	7.75
fare 'go, rush'	13	0.48	12.70	*late* 'pretend'	21	0.77	24.06
svømme 'swim'	12	0.44	13.62	*spørre* 'ask'	21	0.77	3.55
trampe 'stomp'	12	0.44	18.33	*finne* 'find'	19	0.70	0.52
trippe 'walk lightly'	12	0.44	22.66				
tusle 'toddle'	12	0.44	24.15				

(28) Han *brummet* omkring i huset og *pakket* tingene sine (LBK)
he growl.PST around in.the.house and pack.PST the.things his
'He went around the house growling, packing his things'

The top V1 *gå* 'go, walk' was vastly more frequent than the rest, with almost half the tokens and infinite collostructional strength. Rather frequent were also *reise* 'travel', *løpe* 'run', and *vandre* 'wander'. Five more verbs had more than 20 tokens. Two V1s of an overall very high corpus frequency (*være* 'be' and *komme (seg)* 'come' (10 tokens))[13] had negative collostructional strength, i.e. were repulsed by

13. V1 *komme* (*seg*) is more properly regarded as representing two microconstructions: *komme rundt og V* 'come around and V' and *komme seg rundt og V* '(be able to) get around and V'. But in order to be able to calculate collostructional strength, I treated it as one verb.

the construction. But the rest of those with at least 10 tokens were significantly attracted to the construction, with collostructional strengths from 5.41 upwards.

Not only V1s, but also many frequent V2s had significant positive collostructional strength – apart for some repulsed verbs with a high corpus frequency. Among the V2s with at least 10 tokens, the greatest strengths were found for *lete* 'search', *leke* 'play (games)', *snuse* 'sniff, snoop', *rope* 'call, yell', *samle* 'collect', *late* 'pretend', *snakke* 'talk', *se* 'see, look', and *kikke* 'peek'. This gives a certain idea of what one typically 'goes around doing'.

The top nine V1s mentioned were studied further with respect to PF1 constituents and V2 collexemes.[14] As described in Section 3.4, a higher frequency of VroV (with only the atelicizing adverb in PF1) in relation to VrXoV may be an indicator of grammaticalization. The numbers for the nine V1s are given in Table 2. In addition to *gå* 'go, walk', three verbs with the meaning 'run' have markedly higher relative frequencies of VroV than the rest. (*Fly* basically means 'fly', but in this construction it is more frequently used about quick motion on foot.) This may indicate a higher degree of grammaticalization for these four V1s.

Table 2. Frequencies of VroV for the most frequent V1s in the Vr(X)oV subschema

V1	Vr(X)oV	VroV	
	N	N	%
springe 'run'	48	37	77.08
gå 'go, walk'	1355	1002	73.95
løpe 'run'	146	102	69.86
fly 'fly, run'	53	37	69.81
reise 'travel'	217	105	48.39
kjøre 'drive'	57	27	47.37
dra 'go'	46	20	43.48
rusle 'potter'	45	18	40.00
vandre 'wander'	93	36	38.71

Distinctive collexeme V2s were identified for the same nine V1s.[15] V2s were included if, for a given V1, they were represented with at least two tokens and their observed frequency was higher than the expected frequency. The results are presented in Tables 3–5, in order of ascending number of V2 collexemes.

14. This cut-off point was chosen because of the marked distance in frequency between *rusle* 'potter' at rank 9 with 45 tokens and the verbs at rank 10 with 20 tokens.

15. The distinctiveness of collexemes is measured as collostructional strength, see note 12. In this case, the two-by-two contingency table has the following four figures: the frequency of a given V2 after a given V1, the frequency of the V2 after other V1s, the frequency of other V2s after the V1, and the frequency of other V2s after other V1s.

Table 3. Distinctive V2 collexemes of V1s *springe*, *rusle*, and *fly* in the Vr(X)oV subschema

V1: *springe* 'run'			**V1: *rusle* 'potter'**			**V1: *fly* 'fly, run'**		
V2	**N**	**CS**	**V2**	**N**	**CS**	**V2**	**N**	**CS**
leke 'play (games)'	5	3.59	*pirke* 'poke, pick'	2	3.10	*drite* 'shit'	2	2.96
vise 'show'	2	1.56	*titte* 'look, peek'	2	2.28	*bjeffe* 'bark'	2	2.66
			lure 'wonder'	2	1.80	*skrike* 'scream'	3	2.49
						lete 'search'	5	1.58

Table 4. Distinctive V2 collexemes of V1s *kjøre*, *dra*, and *vandre* in the Vr(X)oV subschema

V1: *kjøre* 'drive'			**V1: *dra* 'go'**			**V1: *vandre* 'wander'**		
V2	**N**	**CS**	**V2**	**N**	**CS**	**V2**	**N**	**CS**
lete 'search'	8	3.41	*hente* 'fetch'	2	2.40	*omvende* 'convert'	2	2.94
levere 'deliver'	2	2.90	*tigge* 'beg'	2	2.40	*merke* 'notice'	2	2.47
stjele 'steal'	2	2.39	*jakte* 'hunt'	2	2.26	*høre* 'hear, listen'	2	1.80
hente 'fetch'	2	2.21	*kjøpe* 'buy'	2	2.03	*studere* 'study'	2	1.80
			synge 'sing'	2	1.45	*gruble* 'brood, ponder'	2	1.55
						suge 'suck, inhale'	2	1.45

Some of the V1s had few V2 collexemes. Still, some observations may be made. Most V2 collexemes denote activities that are naturally performed by a subject referent moving from place to place. The collexemes of *fly* 'fly, run' and *løpe* 'run' show a certain tendency to denote noisy or irritating activities. Several of the collexemes of *reise* 'travel' indicate professional performance, including the top V2 *holde* 'hold', which is frequently found as *holde foredrag* 'give a lecture'.

For most V1s except *gå* 'go, walk', there were few collexemes denoting cognition. The exceptions are *lure* 'wonder' for V1 *rusle* 'potter' and *gruble* 'brood, ponder' for V1 *vandre* 'wander'. These V1s denote slow motion on foot, which is typically also the case for *gå*. This verb has a high overall corpus frequency with an immense range of uses and may often be translated with English *go*. Still, its central meaning is that of 'walk'. Many of its distinctive collexemes denote cognition: *tro* 'think, believe', *tenke* 'think, ponder', *huske* 'remember', *vente* 'wait', *drømme* 'dream', *innbille* 'imagine', *vite* 'know'. In addition, *være* 'be' often occurs with adjectives of cognition such as *redd* 'afraid', and *bære* 'carry' is often used in expressions like *bære på en hemmelighet* 'carry a secret'. These cognitive activities or states do not require the subject referent to actually walk around. Certain other collexemes do indicate that *gå* actually refers to walking, e.g. *hilse* 'greet' and *dele*,

Table 5. Distinctive V2 collexemes of V1s *løpe*, *gå*, and *reise* in the Vr(X)oV subschema

V1: *løpe* 'run'			**V1: *gå* 'go, walk'**			**V1: *reise* 'travel'**		
V2	**N**	**CS**	**V2**	**N**	**CS**	**V2**	**N**	**CS**
leke 'play (games)'	17	12.99	*være* 'be'	112	13.70	*holde* 'hold'	20	16.11
rope 'call, yell'	13	8.21	*tro* 'think, believe'	41	11.24	*opptre* 'perform'	4	3.74
dirigere 'direct'	2	2.55	*tenke* 'think, ponder'	42	7.42	*underholde* 'entertain'	3	3.30
bråke 'make noise'	2	2.08	*huske* 'remember'	22	6.72	*samle* 'collect'	9	3.01
mase 'nag, fuss'	2	2.08	*si* 'say'	41	4.05	*besøke* 'visit'	4	2.68
skrike 'scream'	4	2.06	*vente* 'wait'	26	2.78	*tale* 'make speech'	3	2.36
slå 'hit'	4	2.06	*bære* 'carry'	12	2.77	*treffe* 'meet'	3	2.36
hyle 'squeal'	2	1.80	*drømme* 'dream'	8	2.44	*spille* 'play (music, sport)'	5	2.33
sparke 'kick'	2	1.43	*hilse* 'greet'	10	2.23	*undervise* 'teach'	3	1.86
rydde 'tidy'	3	1.37	*smile* 'smile'	14	1.74	*behandle* 'treat'	2	1.75
			kjenne 'feel'	12	1.70	*delta* 'participate'	2	1.75
			innbille 'imagine'	8	1.69	*male* 'paint (art)'	2	1.75
			kalle 'call, name'	10	1.67	*oppleve* 'experience'	2	1.75
			lukte 'smell'	5	1.52	*treske* 'thresh'	2	1.75
			skryte 'brag'	5	1.52	*tilby* 'offer'	2	1.47
			vite 'know'	17	1.39	*utforske* 'explore'	2	1.47
			dele 'hand (out)'	12	1.33	*utføre* 'perform'	2	1.47
						verve 'enlist'	2	1.47
						bli 'become, be'	4	1.31

typically *dele ut* 'hand out', but the preponderance of cognitive V2 collexemes does indicate a fair degree of bleaching of V1 *gå*.

If *gå* is bleached and does not necessarily refer to walking, what meaning or function does it have? This question cannot be answered in detail here. Yet consider the following examples:

(29) Kanskje jeg lurer meg selv der jeg *går* rundt og *tror* at jeg har …
maybe I fool me self where I go.PRS around and think.PRS that I have
'Maybe I'm fooling myself when I go around thinking that I have …' (LBK)

(30) Jeg påstår … ikke at alle norske sosialister *går* rundt og *tenker*
I claim not that all Norwegian socialists go.PRS around and think.PRS
slik (LBK)
thus
'I don't claim that all Norwegian socialists go around thinking like that'

As I interpret these examples compared to the alternatives without *går rundt og*, they emphasize the presence of the belief/thinking during the subject referent's various acts and activities, rather than construing it as a property of the subject referent. This ties in with the characteristic of pseudocoordination that it denotes specific events rather than general properties (see Section 2.4). See further Hesse (2009) for discussion of the aspectual properties of atelic *gå*.

4.2 Second corpus study: VoV

This corpus study targets VoV, a pseudocoordination subschema of atelic motion without any constituents in PF1. The absence of an atelicizing adverb makes the construction formally identical to the pseudocoordination subschema of telic motion without an adverbial expression of direction – compare atelic *gå og smile* 'go (around) smiling' to telic *gå og åpne døra* 'go and open the door' (see Hesse, 2009).

The only substantial part of VoV is the coordinator, and there are two agreeing verbs. This is very little to base a corpus search on. Although PF1 is empty, in matrix clauses there may be constituents in the midfield, esp. subjects and sentence adverbials (see Section 2.1). To include examples with finite verbs, one would need to allow for words intervening between V1 and *og*. Any large-scale corpus search that targeted finite-verb examples and/or an unspecified V1 would have yielded quite intractable amounts of irrelevant hits.

Three search limitations were therefore made. First, no words were allowed between V1 and *og* 'and'. Second, only nonfinite constructions (i.e. with verbs in the infinitive or perfect participle) were searched for, since these will at any rate normally have no verb-phrase external constituents between V1 and *og*. Third, I specified V1 to be one of the verbs that had at least five V1 tokens in the first study (53 different verbs). I judged it unlikely that absent or very low-frequency verbs in the first study would be used in VoV (cf. Josefsson, 1991); without an atelicizing adverbial, most motion verbs strongly favour a telic reading.

The searches yielded 3,646 hits, which were downloaded to a spreadsheet and categorized. Examples of atelic VoV were separated from examples of telic and nonmotion pseudocoordination, and from various other irrelevant hits. There were 372 relevant tokens, two of which are given in (31) and (32):

(31) Jeg har lenge *gått* og *vurdert* om jeg skulle ta permisjon en
I have long go.PTCP and consider.PTCP if I should take leave a
periode (LBK)
period
'I've long been considering whether to take leave for some time'

(32) Fiske og jakte, *rusle* og *se* på den vakre naturen (LBK)
fish and hunt potter.INF and see.INF on the beautiful the.nature
'Fishing and hunting, wandering around looking at the beautiful nature'

The 372 tokens were distributed over nine V1s. Thus, 44 of the 53 V1s searched for were not found. Here *gå* 'go, walk' accounted for more than 90 % of the tokens. Frequency and collostructional strength for the nine V1s are shown in Table 6. The table also shows figures for the most frequent V2s.

Table 6. All V1s and the most frequent V2s in the VoV subschema

V1	N	%	CS	V2	N	%	CS
gå 'go, walk'	337	90.59	∞	*tenke* 'think, ponder'	24	6.43	15.95
reise 'travel'	19	5.11	21.28	*se* 'see, look'	21	5.63	3.07
løpe 'run'	8	2.15	9.33	*vente* 'wait'	18	4.83	18.44
kjøre 'drive'	2	0.54	0.99	*bære* 'carry'	16	4.29	19.43
rusle 'potter'	2	0.54	3.88	*være* 'be'	11	2.95	–2.48
fare 'go, rush'	1	0.27	1.36	*slenge* 'lie/go about'	7	1.88	19.06
fly 'fly, run'	1	0.27	0.90	*glede (seg)* 'be happy'	6	1.61	8.13
sveve 'hover'	1	0.27	1.66	*gruble* 'brood, ponder'	6	1.61	12.17
svinse 'flit'	1	0.27	2.98	*holde* 'hold'	6	1.61	1.45
				håpe 'hope'	6	1.61	7.21

Gå 'go, walk' turned out to have no V2 collexemes that reached significance. Since this V1 dominates so strongly, there is simply little to distinguish it from; its list of most frequent V2s resembles that of the subschema in Table 6. They were: *tenke* 'think, ponder' (24 tokens), *vente* 'wait' (18), *se* 'see, look' (16), *bære* 'carry' (15), *være* 'be' (9), *slenge* 'lie/go about' (7), *glede (seg)* 'be happy, look forward', *gruble* 'brood, ponder', *holde* 'hold', *håpe* 'hope' (6 each), *leie* 'hold hands', *lure* 'wonder', *tråkke* 'plod', *ønske* 'wish' (5 each). For V1 *reise* 'travel', two V2s with at least two tokens were found to be distinctive collexemes: *oppleve* 'experience' (5 tokens,

CS 6.70) and *møte* 'meet' (2 tokens, CS 2.61). V1 *rusle* 'potter' had two tokens, both with V2 *se* 'see, look' (CS 2.52).

4.3 Subschema comparison

In this section, the subschemas Vr(X)oV and VoV are compared. Since only nonfinite tokens of VoV are studied in Section 4.2, only nonfinite tokens of Vr(X)oV are considered in the comparison. Further, I divide the Vr(X)oV subschema into the variants VrXoV and VroV (see Section 4.1).

The relative frequencies of the most frequent verbs in the V1 paradigms are visualized in Figure 1. It is very clear that a reduced formal distance between V1 and V2 is strongly associated with higher relative frequencies for *gå* 'go, walk'.

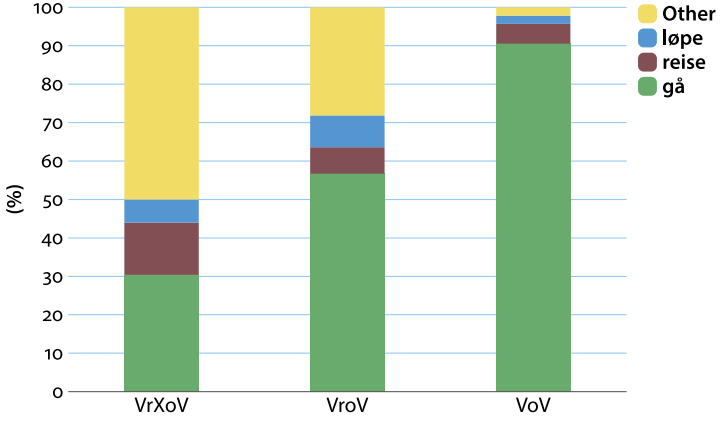


Figure 1. Relative frequencies of V1s in subschemas VrXoV, VroV, and VoV (only nonfinite verbs)

The three subschemas were compared pairwise (VrXoV to VroV and VroV to VoV) in order to identify distinctive collexemes among the V1s and V2s. The results are shown in Tables 7–10. For V1s, *gå* is distinctive of the shorter variant in each compared pair, and its distinctivity is in each case far stronger than that of the distinctive collexemes of the longer variant.

For V2s, one may note the cognition verbs on the lists of distinctive collexemes for the shorter variants in each subschema pair, notably *tro* 'believe' and *huske* 'remember' (and partly *være* 'be', see Section 4.1) for VroV compared to VrXoV, and *vente* 'wait', *tenke* 'think, ponder', *glede (seg)* 'be happy, look forward', *håpe* 'hope', *ønske* 'wish', *lengte* 'long', *sørge* 'grieve' (and maybe *oppleve* 'experience') for VoV.

Table 7. Distinctive V1 collexemes of subschemas VrXoV and VroV

VrXoV			VroV		
V2	**N**	**CS**	**V2**	**N**	**CS**
reise 'travel'	34	2.49	gå 'go, walk'	240	10.69
rusle 'potter'	12	2.32			
komme (seg) 'come, get'	5	1.54			
dra 'go'	9	1.46			

Table 8. Distinctive V1 collexemes of subschemas VroV and VoV

VroV			VoV		
V2	**N**	**CS**	**V2**	**N**	**CS**
løpe 'run'	35	4.12	*gå* 'go, walk'	337	27.76
vandre 'wander'	10	2.76			
fly 'fly, run'	11	2.25			
kjøre 'drive'	10	1.51			
dra 'go'	5	1.37			
springe 'run'	5	1.37			

Table 9. Distinctive V2 collexemes of subschemas VrXoV and VroV

VrXoV			VroV		
V2	**N**	**CS**	**V2**	**N**	**CS**
samle 'collect'	6	1.92	*være* 'be'	37	2.19
rope 'call, yell'	8	1.76	*tro* 'believe'	10	2.03
			huske 'remember'	8	1.62
			lete 'search'	14	1.53
			gjøre 'do'	10	1.36

5. Concluding discussion

Against the background of a theoretical discussion, the present study has explored the properties of a set of subschemas of Norwegian pseudocoordination, viz. those of atelic motion, using a corpus of contemporary written Norwegian Bokmål of about 100,000,000 words. It has repeatedly been claimed in the literature (see Section 3.3) that pseudocoordination allows only a limited number of first verbs – indicating paradigmatization and some degree of grammaticalization. However, the first corpus study in Section 4 finds 167 different V1s used in pseudocoordina-

Table 10. Distinctive V2 collexemes of subschemas VroV and VoV

VroV			VoV		
V2	**N**	**CS**	**VoV**	**N**	**CS**
være 'be'	37	3.40	*vente* 'wait'	18	6.03
lete 'search'	14	1.92	*bære* 'carry'	16	3.66
gjøre 'do'	10	1.51	*tenke* 'think, ponder'	24	3.40
si 'say'	13	1.37	*glede (seg)* 'be happy'	6	1.99
			håpe 'hope'	6	1.99
			oppleve 'experience'	5	1.66
			ønske 'wish'	5	1.66
			slenge 'lie/go about'	7	1.64
			drive 'go about'	4	1.32
			lengte 'long'	4	1.32
			sørge 'grieve'	4	1.32

tion of atelic motion. While other subschemas may exhibit fewer V1s, this result shows that the V1 paradigm of the pseudocoordination schema is far from small.

Still, closer examination of the atelic-motion subschema and its subsidiary subschemas reveals that some V1s are more central than others. In particular, *gå* 'go, walk' is by far the most frequent V1, and its dominance in terms of relative frequency strengthens as the formal distance between it and V2 diminishes. Decreasing distance above all means that there are fewer locative adverbials in PF1. Such use is uncommon for motion verbs outside of pseudocoordination, and it indicates that, in pseudocoordination, esp. *gå* (but also certain verbs meaning 'run') tends to have a reduced argument structure in this construction, as found by Hilpert & Koops (2008). It seems warranted to regard this as an instance of decategorialization, which in turn may be interpreted as an indication of a certain degree of grammaticalization.

The decreasing superficial distance between V1 and *og* could be taken to indicate condensation (narrowing of structural scope) or coalescence (increasing bondedness), cf. Lehmann (2015). However, there does not seem to be any clear evidence that the constructional variants of atelic motion with less material intervening between V1 and *og* have a different morphosyntactic structure from those with more material (apart from the absence of that particular material, of course); VoV, VroV and VrXoV may all be analyzed as verb phrase coordinations. Even in VoV there may be intervening constituents in matrix clauses (see Section 2.1), viz. subjects and sentence adverbials, which presumably disfavour and slow down processes of increasing morphosyntactic integration. However, the inability of a verb like *drive* 'carry on' to be used in pseudocoordination with locative adverbials

in PF1 (see Section 3.4) suggests that constructions with this and certain other V1s involve clearer changes in morphosyntax. This issue cannot be resolved here.

Further, distinctive collexeme analysis has uncovered that esp. *gå* exhibits a stronger tendency than other V1s to collocate with cognitive V2s. This tendency points to V1 bleaching; the motion meaning is often virtually absent (cf. Hesse, 2009). The semantic or functional result of this bleaching is rather hard to describe (see Section 4.1), but it may certainly be considered an indicator of a notable degree of grammaticalization.

Thus, one V1 that is strongly associated with the construction also exhibits a higher degree of grammaticalization both syntactically and semantically. In Bybee's (2010, p. 90) words, "diachronic development seems to emanate outwardly from the central member of a category". It cannot be determined on the basis of synchronic data whether *gå* leads on in a development followed by other frequent V1s like e.g. *løpe* 'run', but the verb certainly has a central role in the synchronic category structure.

Except for the coordinator, no slot in the schema-level construction of pseudocoordination is always host to a grammaticalized element. Thus, beyond the coordinator, pseudocoordination does not necessarily entail grammaticalization. But when V1 is grammaticalized, other parts of the construction and the relations between them are affected in various ways (cf. Himmelmann (2004) and Traugott & Trousdale (2013)). The distance between V1 and V2 tends to be short due to the absence of locative adverbials. To the extent that the constructs are of the maximally dense VoV type and the meaning contribution of V1 is primarily aspectual or something similar, the construction resembles auxiliary verb constructions. The facilitation relation that holds between VP1 and VP2 in typical pseudocoordination is more blurred in such cases due to the primarily grammatical semantics of V1. The bleaching of V1 also affects the meaning of the construction as a whole, since VP2 contributes most of the lexical meaning, and this strengthens the single-event meaning that already characterizes pseudocoordination in general. Even the coordinator is affected, since its additive meaning is obscured when one of its participants is strongly bleached.

References

Ahlberg, M., Andersson, P., Forsberg, M., & Tahmasebi, N. (2015). A case study on supervised classification of Swedish pseudo-coordination. *20th Nordic Conference of Computational Linguistics (NODALIDA 2015), NEALT Proceedings Series*, 23, 11–19.

Aikhenvald, A. (2006). Serial verb constructions in typological perspective. In A. Aikhenvald & R.M.W. Dixon (Eds.), *Serial Verb Constructions* (pp. 1–68). Oxford: Oxford University Press.

Bachmann, I. (2013). Has *go-V* ousted *go-and-V*? In H. Hasselgård, J. Ebeling, & S.O. Ebeling (Eds.), *Corpus Perspectives on Patterns of Lexis* (pp. 91–111). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.57.09bac

Barðdal, J. (2008). *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8

Bjerre, T., & Bjerre, A. (2007). Pseudocoordination in Danish. In S. Müller (Ed.), *Proceedings of the HPSG07 Conference* (pp. 6–24). Stanford, CA: CSLI Publications.

Blensenius, K. (2015). *Progressive Constructions in Swedish*. Dissertation. Gothenburg: Göteborgs Universitet.

Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language*, 88, 1–44.

Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Chaves, R.P. (2007). *Coordinate Structures. Constraint-Based Syntax-Semantics Processing*. Dissertation. Lisbon: Universidade de Lisboa.

Faarlund, J.T., Lie, S., & Vannebo, K.I. (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.

Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.

Gries, S.Th. (2014). *Coll.analysis 3.5. A script for R to compute perform collostructional analyses*. http://tinyurl.com/collostructions

Hansen, E., & Heltoft, L. (2011). *Grammatik over det Danske Sprog*. Copenhagen: Det Danske Sprog- og Litteraturselskab.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization*. Chicago: The University of Chicago Press.

Heine, B., & Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511613463

Hesse, A. (2009). *Zur Grammatikalisierung der Pseudokoordination*. Tübingen: A. Francke.

Heycock, C., & Petersen, H.P. (2012). Pseudo-coordinations in Faroese. In K. Braunmüller & G. Christoph (Eds.), *Multilingual Individuals and Multilingual Societies* (pp. 259–280). Amsterdam: John Benjamins. https://doi.org/10.1075/hsm.13.18hey

Hilpert, M. (2008). *Germanic Future Constructions*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.7

Hilpert, M. (2010). What can synchronic gradience tell us about reanalysis? In E.C. Traugott & G. Trousdale (Eds.), *Gradience, Gradualness and Grammaticalization* (pp. 181–201). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.90.10hil

Hilpert, M., & Koops, C. (2008). A quantitative approach to the development of complex predicates. *Diachronica*, 25, 240–259. https://doi.org/10.1075/dia.25.2.06hil

Himmelmann, N. P. (2004). Lexicalization or grammaticalization: Opposite or orthogonal? In W. Bisang, N. P. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization?* (pp. 21–42). Berlin: De Gruyter.

Höder, S. (2012). Der is wieder bei und malt Karten. *Niederdeutsche Syntax aus nordeuropäischer Sicht*. In R. Langhanke, K. Berg, M. Elmentaler, & J. Peters (Eds.), *Niederdeutsche Syntax* (pp. 181–201). Hildesheim: Georg Olms.

Holm, E. (2013). *"Det sto aldri noe på papir, men det lå i luften og satt i veggene". Kroppspositurverb brukt om eksistens og lokasjon i norsk*. MA thesis. Oslo: Universitetet i Oslo.

Hopper, P. J. (1991). On some principles of grammaticization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*, Volume 1 (pp. 17–36). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.19.1.04hop

Hopper, P. J. (2002). Hendiadys and auxiliation in English. In J. Bybee & M. Noonan (Eds.), *Complex Sentences in Grammar and Discourse* (pp. 145–173). Amsterdam: John Benjamins. https://doi.org/10.1075/z.110.09hop

Huldén, L. (1961). Om samordning hos Bellman. *Studier i nordisk filologi*, 51, 7–149.

Johannessen, J. B. (1998). *Coordination*. Oxford: Oxford University Press.

Johnsen, L. G. (1988). A note on subcoordination. *Trondheim Working Papers in Linguistics*, 6, 195–201.

Josefsson, G. (1991). Pseudocoordination – a VP + VP coordination. *Working Papers in Scandinavian Syntax*, 47, 130–56.

Kinn, T. (2014). Verbalt presens partisipp. *Norsk Lingvistisk Tidsskrift*, 32, 62–99.

Kinn, T. (2017). Den trassige verbforma: presens partisipp og kongruensproblem i pseudokoordinasjon. In Z. S. Hansen, A. Johansen, H. P. Petersen, & L. Reinert (Eds.), *Bók Jógvan* (pp. 213–230). Tórshavn: Fróðskapur. Faroe University Press.

Kjeldahl, A. (2010). *The Syntax of Quirky Morphology*. Dissertation. Aarhus: Aarhus Universitet.

Kuteva, T. (1999). On 'sit'/'stand'/'lie' auxiliation. *Linguistics*, 37, 191–213. https://doi.org/10.1515/ling.37.2.191

Kuteva, T. (2001). *Auxiliation*. Oxford: Oxford University Press.

Kvist Darnell, U. (2008). *Pseudosamordningar i svenska*. Dissertation. Stockholm: Stockholms universitet.

Lakoff, G. (1986). Frame semantic control of the coordinate structure constraint. *Chicago Linguistic Society*, 22, 152–67.

Langacker. R. W. (1991). *Foundations of Cognitive Grammar, Volume II: Descriptive Application*. Stanford, CA: Stanford University Press.

Lehmann, C. (2015). *Thoughts on Grammaticalization*. 3rd edition. Berlin: Language Science Press.

Lemmens, M. (2005). Aspectual posture verb constructions in Dutch. *Journal of Germanic Linguistics*, 17, 183–217. https://doi.org/10.1017/S1470542705000073

Lesuisse, M., & Lemmens, M. (this volume). Grammaticalisation cut short: a diachronic constructional view on English posture verbs.

Lødrup, H. (2002). The syntactic structures of Norwegian pseudocoordinations. *Studia Linguistica*, 56, 121–43. https://doi.org/10.1111/1467-9582.00090

Lødrup, H. (2014). There is no reanalysis in Norwegian pseudocoordinations (except when there is). In H. P. Helland & C. M. Salvesen (Eds.), *Affaire(s) de grammaire* (pp. 43–65). Oslo: Novus.

Newman, J., & Rice, S. (2004). Patterns of usage for English *sit* , *stand*, and *lie* . *Cognitive Linguistics*, 15, 351–96. https://doi.org/10.1515/cogl.2004.013

Newman, J., & Rice, S. (2008). Asymmetry in English multi-verb sequences. In B. Lewandowska-Tomaszczyk (Ed.), *Asymmetric Events* (pp. 3–23). Amsterdam: John Benjamins. https://doi.org/10.1075/celcr.11.03new

Nielsen, P. J. (2011). *Kongruenskonstruktion i dansk*. Oslo: Novus.

Ross, J. R. (1967). *Constraints on Variables in Syntax*. Dissertation. Cambridge, MA: Massachusetts Institute of Technology.

Stefanowitsch, A. (2000). The English GO-(PRT)-AND-VERB construction. *Berkeley Linguistics Society*, 26, 259–70.

Stefanowitsch, A. (2013). Collostructional analysis. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 290–306). Oxford: Oxford University Press.

Stefanowitsch, A., & Gries, S. Th. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8, 209–43. https://doi.org/10.1075/ijcl.8.2.03ste

Teleman, U., Hellberg, S., & Andersson, E. (1999). *Svenska Akademiens grammatik*. Stockholm: Svenska Akademien.

Thorell, O. (1973). *Svensk grammatik*. Stockholm: Esselte Studium.

Tonne, I. (1999). A Norwegian progressive marker and the level of grammaticalization. *Languages in Contrast*, 2, 131–59. https://doi.org/10.1075/lic.2.1.08ton

Tonne, I. (2001). *Progressives in Norwegian and the Theory of Aspectuality*. Dissertation. Oslo: Universitetet i Oslo.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Trousdale, G. (2014). On the relationship between grammaticalization and constructionalization. *Folia Linguistica*, 48, 557–77. https://doi.org/10.1515/flin.2014.018

Vannebo, K. I. (1969). *Aksjonsart i norsk*. Oslo: Universitetsforlaget.

Vannebo, K. I. (2003). *Ta og ro deg ned noen hakk*: On pseudocoordination with the verb *ta* 'take' in a grammaticalization perspective. *Nordic Journal of Linguistics*, 26, 165–93. https://doi.org/10.1017/S0332586503001070

de Vos, M. A. (2005). *The Syntax of Verbal Pseudo-coordination in English and Afrikaans*. Dissertation. Leiden: Universiteit Leiden.

Wierzbicka, A. (1980). *Lingua mentalis: The Semantics of Natural Language*. London: Academic Press.

Wiklund, A.-L. (2007). *The Syntax of Tenselessness*. Berlin: De Gruyter.

Wulff, S. (2006). *Go-V* vs. *go-and-V* in English: A case of constructional synonymy? In S. Th. Gries & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics* (pp. 101–125). Berlin: De Gruyter.

Evaluating grammaticalization and constructional accounts

The development of the inchoative construction with *put* verbs in Spanish

Renata Enghels & Marie Comer
Ghent University

This article explores the historical development of the inchoative construction in Spanish with the *put* verbs *poner* and *meter* filling the auxiliary verb slot. Previous studies have focused mainly on the syntactic and semantic processes through which full lexical *put* verbs come to fulfil an auxiliary role in the inchoative periphrasis. The analysis presented in this article aims at going beyond this traditional, grammaticalization account, and examines whether a constructional approach offers a more nuanced account of the empirical data. The main objective is to verify the extent to which different structural patterns of the inchoative construction have given rise to a general constructional schema.

Keywords: inchoative construction, *put* verbs, Spanish, grammaticalization, diachronic constructional studies

1. Introduction

In order to express inchoative aspect, Spanish speakers dispose of a wide gamut of verbal periphrasis. Different lexical verb classes are used as auxiliary verb for this purpose. They include lexically ingressive verbs (*empezar a/comenzar a bailar* 'to start to dance'), change of state verbs (*romper a llorar*, lit. 'to break to cry', 'to start to cry'), movement verbs (*echarse a reír*, lit. 'to throw oneself to laugh', 'to start to laugh') and *put* verbs (*ponerse/meterse a escribir*, lit. 'to put oneself to write', 'to start to write') (Gómez Manzano, 1992; Gómez Torrego, 1999; García Fernández et al., 2006; Aparicio et al., 2014). This article focuses on the development of the Spanish inchoative construction using the last of these categories, and its relationship with the intricate semantics and argument structure of its verbal nucleus.

https://doi.org/10.1075/cal.21.c5

The process by which the above-listed full lexical verbs come to fulfil an auxiliary role in the inchoative periphrasis is particularly suited to study within grammaticalization theories (e.g. Lamiroy & De Mulder, 2011). Adopting a mainly semasiological perspective, these studies ask how an individual referential expression acquires procedural meaning by means of processes of syntactic reduction and semantic bleaching. Thus Verroens (2011) reconstructs the grammaticalization process which has allowed the French *put* verb *mettre* to function as an auxiliary verb in an inchoative periphrasis. This reconstruction is an expansion of Heine's (2002) four-step grammaticalization model,[1] and proposes a six-stage scenario:

1. In an initial stage, *mettre* occurs with its locative source meaning (an agent displaces a theme to another location in space), and has the following argument structure [NP $_1$ + V *mettre* + NP $_2$ + LOC] (e.g. *je mets le livre dans ma chambre*, 'I put the book in my room').
2. In an intermediate stage, the theme can be an animate entity without control or willfulness in being displaced (e.g. *je mets Jean dans ma chambre*, 'I put John in my room').
3. The bridging context gives rise to a certain ambiguity: the agent causes a theme to move from one place to another, but at the same time, s/he makes him/her start an activity or an event (expressed by the infinitive). Syntactically, *mettre* simultaneously combines with a locative complement and an infinitive [NP $_1$ + V *mettre* + NP $_2$ + LOC + (a + INF)] (e.g. *je mets Jean dans ma chambre à faire ses devoirs*, 'I put John in my room to do his homework').
4. In the next stage, the full lexical locative meaning is gradually substituted by a procedural causative one (an agent incites a theme to do something). The verb then takes the following argument structure [NP $_1$ + V *mettre* + NP $_2$ + (a + INF)] (e.g. *Je mets Jean à faire ses devoirs*, 'I set John to do his homework')
5. In the switch context, the verb is used reflexively [NP $_1$ + V *se mettre* + (a + INF)]. The clitic *se* undergoes reanalysis: it loses its autonomy as an independent argument and becomes an integral part of the verb. This corresponds to the inchoative reading (e.g. *Jean se met à faire ses devoirs* 'John starts doing his homework').

1. As is well known, the four stages are: (1) the initial stage, in which a linguistic expression has its (original, non grammaticalized) source meaning; (2) the bridging contexts, which are ambiguous between the source and target meaning: rather than the source meaning, a new (target) meaning offers a more plausible interpretation of the utterance concerned; (3) the switch contexts, in which the target meaning of the linguistic expression offers the only possible interpretation: the source meaning is ruled out and ambiguous readings disappear; (4) the stage of conventionalization, including extensions toward new contexts of use.

6. In a final stage the construction *se mettre à + inf* conventionalizes and spreads throughout the linguistic community.

Although this model offers a plausible explanation for the grammaticalization of a lexical *put* verb into an aspectual marker, it remains as yet a hypothetical design which has not been empirically verified on the basis of diachronic data. Moreover, the model has a number of empirical and theoretical implications which raise doubts about its applicability to cognate Spanish inchoatives.

First, whereas the French inchoative periphrasis is restricted to the *put* verb *mettre*, the Spanish equivalent can be construed with not one, but two *put* verbs (*poner* and *meter*) (Aparicio et al., 2014). Given this difference, it needs to be examined whether generalizations can be made over the historical development of particular instances, and thus whether the inchoative periphrasis with *poner* has developed in a similar way to that with *meter*, and by extension to inchoative periphrases with other verbal nuclei.

Second, the process of reanalysis is presented as a straightforward unidirectional model which takes only tangential account of the availability of different linguistic choices to the speaker. In fact, historically, the Spanish inchoative periphrasis included not only the preposition *a* but also *en* (e.g. *se metieron a/en comprar una casa* 'they started to look into buying a house'), and different word orders with or without intercalation of a lexical element between the *put* verb and the subordinated infinitive are attested (e.g. *se pone a leer en su cuarto* vs. *se mete en su cuarto a leer*, 'he starts to read in his room'). To put this another way, grammaticalization models set off from the idea that changes pertain to an individual linguistic item, and do not always pay full attention to the features of the immediate contexts in which the item is used.

This paper explores the extent to which these two drawbacks can be overcome by integrating findings from grammaticalization research into the framework of construction grammar. Instead of studying the development of individual lexical items, diachronic constructional studies explore the history of syntactic patterns, including how new constructions come into being and gradually develop new functions (Noël, 2007; Bergs & Diewald, 2008; Hilpert, 2008, 2013; Trousdale, 2008; Traugott & Trousdale 2013; Coussé 2018; among others).

The central idea of the constructional approach is that different expressions (such as different inchoative periphrases) with analogous meanings and formal properties are instances of a more general abstract schematic structure or construction. The language system is defined as "a network of conventionalized pairings of form and meaning, in which specific micro-constructions inherit properties from more general schemas" (Trousdale, 2014, p. 557). Meaning is carried by constructions as a whole, and abstracts away from the input of individual lexical

items. In this approach, the paradigm of inchoative periphrases can be defined as a conventional form/meaning pairing for which the use of the term 'construction' seems appropriate (Goldberg, 1995; Croft, 2001). Formally, both reflexive verbs *ponerse* and *meterse* (as well as other verbs such as *comenzar*, *empezar*, *echarse* etc.) function as phonologically substantive elements and occupy a fixed auxiliary slot [V]. They are accompanied by three schematic positions or 'open' slots: a subject [NP $_1$], a preposition [PREP] and an infinitive slot [INF]. Semantically the subject [NP $_1$] of the inchoative construction forces or incites him/her or itself to carry out the event expressed by the infinitive complement in (1).

(1) [NP $_1$ + V $_{refl}$ + (PREP + INF)] ↔ agent starts the activity Z
   a. *ponerse a bailar*, 'to start to dance'
   b. *meterse a dormir*, 'to fall asleep'
   c. *comenzar a trabajar*, 'to start to work'; *echarse a reír*, 'to start to laugh'; etc.

The constructional approach relies on the idea that the taxonomic network consists of inheritance relations between constructions defined at different levels of schematicity. Ranging from the highest degree of abstraction to the highest degree of concreteness, linguistic schemas are instantiated by subschemas, and at lower levels by micro-constructions and constructs (that is, the empirically attested tokens) (Goldberg, 1995; Traugott & Trousdale, 2013).[2] Properties of lower level nodes can be inherited from a superordinate schema, or can be specific to that node. Clearly, the inchoative construction is sanctioned by the higher level caused-motion schema proposed by Goldberg (1995, p. 52) and shown in Figure 1, and more concretely its instantiation through the caused-motion *put* construction. According to this basic schema, an agent (SUBJ) causes a theme (OBJ) to move to another place (namely, the goal, expressed by a locative complement (OBL)).

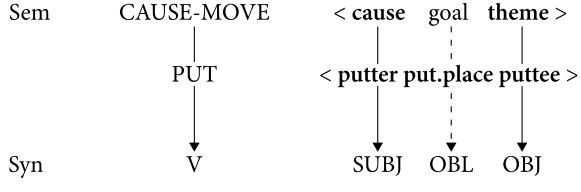


Figure 1. Caused-motion schema applied to *put* (Goldberg, 1995, p. 52)

2. The term micro-construction is generally used whenever a distinction is made between higher-level and lower-level schemas. However, given the current debate on how to define the right level of abstraction at which particular constructions are to be situated (see Hilpert, 2013; Traugott, personal communication), we revert to the more neutral term of 'structural variants' to refer to different instantiations of a more general schema.

Indeed, as prototypical *put* verbs, *poner* and *meter* basically express a caused-motion event involving an object or theme's change of location to a new location in space. Syntactically, this event is rendered by a four-slot locative construction (2) (Cifuentes Honrubia, 1999; Ibarretxe, 2012).

(2) [NP $_1$ + V + NP $_2$ + LOCATIVE COMPL] ↔ agent causes patient to move to a new location in space
   a. *poner el mantel en la mesa*, 'to put the table cloth on the table'
   b. *meterse un chupete en la boca*, 'to put a pacifier in one's mouth'

As has been shown in a previous study (Comer & Enghels, in press), this caused-motion *put* construction can be considered the historical source construction of the inchoative construction. Its immediate parent is to be identified as the 'periphrastic verb construction', combining a conjugated auxiliary verb form with a non-finite (mostly infinitive) form. This construction expresses several aspectual (such as the beginning, ending or development of the event) and modal (such as obligation or necessity) meanings.[3]

Returning to the aforementioned drawbacks of the grammaticalization account, a diachronic constructional approach can, first, allow us to formulate conclusions on both the functional and formal development of the inchoative construction at a more general level, abstracting away from the specific import of individual lexical items. Secondly, a particularly appealing implication of the constructional view is the recognition of relationships between patterns and structural variants within a paradigm, so that the usage of one particular structural variant cannot be seen as independent of other related variants.

Against this background, this article aims to empirically evaluate whether, by applying the constructional approach, we are indeed capable of offering a more nuanced model of the development of inchoative periphrases with *put* verbs. This goal has significance for one of the main research questions running through

3. A more detailed definition of the (Spanish) periphrasis construction falls beyond the scope of this article (see for instance Garachana, 2016 for a consideration of its most important properties). Moreover, note that in its current use, at the subschema level, the inchoative construction relates to the causative construction, in which a subject [NP $_1$] semantically 'forces' or 'incites' the object [NP $_2$] to perform the event expressed by the infinitive in the [INF] slot: [NP $_1$ + V + NP $_2$ + (PREP + INF)] ↔ agents causes patient to start the activity Z: *poner a alguien a bailar* 'to make someone dance (lit. 'put someone to dance'); *meter a alguien a dormir* 'to make someone sleep (lit. 'put someone to sleep'). This network dimension offers valuable insights into the development of the inchoative construction. The most important of these are explicitly pointed to throughout the article. However, a more thorough analysis of the links between the inchoative and causative construction on the one hand, and the more general-caused motion schema on the other, is left for future research.

Hilpert's (2013) case studies on constructional change, namely to what extent have different structural patterns given rise to a general constructional schema?[4] Do patterns dissimilate over time or do they undergo functional and structural leveling? The different patterns would then include different auxiliary verbs (*poner* vs. *meter*), or variable structural properties (such as preposition selection or variable word order – see Section 2.2). As claimed by Hilpert (2013, p. 157) "a leveling process would indicate that an abstract constructional schema is forming, allowing speakers to produce new tokens that mix and match characteristics from different source constructions." This process would then point towards analogical matching, also identified as an important mechanism of change in the constructionalization model (see Traugott & Trousdale, 2013, p. 21 and footnote 4). If, on the contrary, growing divergence is observed, this would point towards a looser connection between individual variants of constructions, and a lack of schema-forming. In line with these considerations, this article pursues a double goal:

1. On a more theoretical level, it assesses the application of a constructional approach to inchoative periphrases (with *put* verbs). Is the inchoative 'subschema' a valid concept, or is the inchoative meaning mainly due to the grammaticalization of an individual lexical item?
2. As concerns the concrete case study, it offers insights into the linguistic parameters and contextual factors which have promoted the development of the inchoative periphrasis in Spanish.

The remainder of this paper is organized as follows. Section 2 provides details of the data collection, and sets out the parameters that define the structural variation within the inchoative construction. Section 3 gives the results of the corpus analysis, and zooms in on the development of the relative token frequency of the inchoative construction with *poner* and *meter* (3.1). In addition, it discusses the evolution of its structural variants defined in terms of the selected preposition (3.2), the internal word order (3.3), and the different semantic types that can be expressed by the inchoative construction (3.4). Finally, Section 4 compares the

4. In this article, the term 'constructional change' is used instead of 'constructionalization'. Contrary to the definition provided by Traugott & Trousdale (2013, p. 22), according to which constructionalization implies the creation of a $form_{new}$/$meaning_{new}$ pairing as a new node in the network, this study is not involved with the creation of the inchoative construction as a new construction. This phase is to be situated in earlier centuries, but is in fact hard to trace (see Section 3.1). What we are observing is a series of "post-constructionalization constructional changes" (in terms of Traugott & Trousdale, 2013, p. 27), abbreviated as 'constructional changes'. As defined by Hilpert (2013, p. 16) this can refer to any change with regard to the function, form, frequency, or distribution of the construction. For a critical assessment of the concept of 'constructionalization', see also the article by Hilpert presented in this volume.

benefits of the constructional approach with those of existing grammaticalization accounts.

2. Materials and method

2.1 Data collection

In order to achieve the above-mentioned goals, this study outlines the development of different patterns within the inchoative construction family between the 13th and 21st century. The analysis relies on a large historical sample of the two Spanish *put* verbs. This sample was compiled from the two reference corpora designed by the Real Academia Española: CORDE (*Corpus Diacrónico del Español*) and CORPES XXI (Corpus of the 21st century). From these were extracted all inchoative periphrases of *poner* and *meter* as defined in (1) above from peninsular Spanish only, from the 13th, 15th, 17th, 19th, and 21st centuries.[5]

The compiled corpus chronologically covers each of the five phases which traditionally divide the history of Spanish language: Medieval Spanish (13th, 15th c.), Classical Spanish (17th c.), Modern Spanish (19th c.) and Contemporary Spanish (21st c.) (Cano Aguilar, 1992). This method of investigating every other century instead of consecutive ones is commonly applied in other diachronic studies of Spanish (see for instance Torres Cacoullos, 2000; Bybee & Torres Cacoullos, 2009; Vergara Wilson, 2009, p. 278, p. 286; Jansegers, 2015, pp. 219–220), and has the particular advantage of revealing important tendencies in the data by covering longer spans of time (Vergara Wilson, 2009, p. 286). Another, more practical, advantage is that it allows us to cover larger time spans with quantitatively more restricted data sets.[6] In total, 4765 tokens were retrieved of the inchoative construction with *poner*, and 378 with *meter* (see Table 1). The inchoative construction with *poner* thus clearly outnumbers that with *meter* (see Section 3.1).

5. CORDE is a large reference corpus containing about 125,000,000 words in documents published between the beginning of Spanish literary production in the Middle Ages (13th c.) and 1975. All documents are written texts, and cover both fiction (prose and poetry) and non-fiction. CORPES XXI contains about 225,000,000 words in documents dated between 2001 and 2016: both fiction and non-fiction, and both written and (to a minor degree) oral texts are included. Given that CORDE is a non-lemmatized corpus, the data were collected by searching for specific verb forms, namely all regular and irregular verb stems of *poner* and *meter*, taking into account the modern and old spelling variants of the verbs. CORPES XXI is lemmatized, which facilitated exhaustive collecting of the data.

6. For more detailed information on the advantages of distinguishing different chronological cross-sections, see Jansegers (2015, pp. 219–220) and Vergara Wilson (2009, p. 278, p. 286).

Table 1. Absolute frequencies of the inchoative construction per time period[7]

	inchoative$_{poner}$	inchoative$_{meter}$	Total
13th c.	7	43	50
15th c.	115	33	148
17th c.	648	86	734
19th c.	1066	144	1210
21st c.	2929	72	3001
Total	4765	378	5143

2.2 Parameters of constructional variation in inchoative periphrases

Constructions are defined as "multifaceted linguistic units that may display simultaneous change at different levels of description – frequency, structure, meaning, and distribution in the linguistic community" (Hilpert, 2013, p. 21). Consequently, in order to comprehend the development of the inchoative construction, we need to trace relevant changes in terms of its form, function, frequency of use, and distribution. Therefore, the starting point for the empirical analysis is that, in order to understand how a construction has developed, it is crucial to observe possible changes in the distribution of different structural and functional variables within a more general schema. These observations can point towards changes in the configuration of a network, such as the creation or decline of particular structures, or an internal reorganization of a network.

Concretely, the inchoative construction comprises a set of constructional variants (or "allostructions", in Cappelle's 2006 terms) which can be defined along four dimensions: (1) the semantics of the *put* verb (*poner* vs. *meter*); (2) the semantics of the preposition (*a* vs. *en*); (3) the internal word order, in particular the placement of adverbial complements; (4) the semantic event type expressed by the construction. Each of these four variables is discussed below.

1. **Lexical verb: *poner* vs. *meter*.** Throughout the analysis, extensive attention is paid to the effect of the lexical source verbs *poner* and *meter* on the evolution of individual micro-constructions. The two verbs differ considerably with regard to their etymological source. Whereas *poner* derives from Latin *pōnĕre*, meaning 'to place', *meter* comes from Latin *mĭttĕre* meaning 'to send, to throw'. In their etymological dictionary of Spanish, Corominas & Pascual (1980–1991) mention that in the Middle Ages, the semantic field of *meter*

7. Unfortunately, the CORDE and CORPES XXI corpora do not provide information on the corpus size for each time slot. Hence, the relative frequencies of the different samples versus the whole corpus cannot be calculated.

extended to encroach on that of locative *poner*. In present-day Spanish, the two verbs are near-synonymous but have semantically specialized in different domains, even if, at first glance, they share a similar locative meaning (Cifuentes, 1999; Comer et al., 2015). The verb *poner* is far more widely used in discourse than *meter*, and is the less semantically specific. It refers to placements both on a horizontal (e.g., *poner/*meter los platos sobre la mesa*, 'to put the plates on the table') and vertical surface (e.g., *poner/*meter el cuadro en la pared*, 'to put the painting on the wall'). The verb *meter*, on the other hand, is preferred for internally directed (concrete or abstract) movements (e.g., *meterse/*ponerse en la vida de alguien*, lit. 'to put oneself in somebody's life). Moreover, compared to *meter, poner* extends more frequently toward other, more grammaticalized domains (such as the pseudo-copulative, e.g. *ponerse nervioso* 'to get nervous'). Consequently, the following research question permeates the empirical analysis: do particular structural variables occur more frequently with one of the two lexical verbs, or does the variability abstract away from the unique semantic profile of the verb and is thus situated at a more abstract constructional level?

2. **Variation in the prepositional slot.** A second source of structural variability is the prepositional slot, which can be filled not only by the preposition *a* ('to') but also by the preposition with locative meaning *en* ('in'). To what extent do the two prepositions alternate freely throughout the history of the inchoative construction, and do we observe a gradual fixation of form of the construction? (Section 3.2)
3. **Internal word order and placement of adverbials.** This variable involves the relative word order within the inchoative construction and, more concretely, the intercalation of lexical elements between the *put* verb slot and the INF slot. Two structural variants are to be distinguished: an adverbial complement can be placed in between the *put* verb and the infinitive (NP $_1$ + V + ADV + (PREP + INF)), e.g., *ponerse en su cuarto a leer* '(lit.) to put oneself in one's room to read') or after the infinitive (NP $_1$ + V + (PREP + INF) + ADV) e.g., *ponerse a leer en su cuarto* 'to start to read in one's room'). This variable relates to the degree of syntactic incorporation and unithood of the inchoative construction, which is, of course, higher when no lexical element appears between the two verbs. It also links up with the degree of compositionality of the construction, defined by Traugott & Trousdale (2013) as the relationship between the meaning of individual items and the meaning of the construction as a whole. A higher degree of unithood within a construction coincides with a reduced degree of compositionality, as the import of the meaning of individual elements becomes less transparent (Section 3.3).[8]

4. **Semantic types.** The fourth and final variable relates to the semantic development of the inchoative construction. Different variants are distinguished in terms of the event type expressed by the construction, for instance an animate participant commencing an activity (*se pone a leer*, 'he starts to read') versus the commencement of a non-volitional event (*se pone a llover*, 'it starts to rain'). As opposed to the formal features explained in (2) and (3) above, where the development of the inchoative construction is expected to go hand in hand with reduced structural variability, this semantic parameter is expected to point towards more variation as the inchoative construction develops. Indeed, a higher degree of constructional vitality coincides with a higher probability of the constructional schema generating new instances (Hilpert, 2013, pp. 116ff; Traugott & Trousdale, 2013; Trousdale, 2014). This can be seen as a correlate of an increasing degree of schematicity of a construction. When constructions are used more frequently, speakers generalize over the individual types and develop more abstract schemas. In its turn, this abstract schema attracts new types, which again contributes to the productivity at the schema level (Barðdal, 2008) (Section 3.4).

3. Results and discussion

To recap: on various accounts, it has been argued that the availability of different options for the speaker is an important locus of language change, and that the changing distribution of constructional variants provides valuable insights into this process (Cappelle, 2006; Hilpert, 2013). Likewise, with regard to the inchoative construction, the language user's preference in the choice between structural and semantic variants is not expected to have remained constant over time. This section examines, on a quantitative and qualitative basis, the extent to which such diachronic development of the relative frequencies of structural and semantic variants can be observed, and whether some general tendencies can be detected.

8. Note that the different structural variants, with and without intercalation of a lexical item, relate to various degrees of reanalysis that characterize the process of emergence of periphrasis verb constructions. All examples included in the corpus express inchoative (and not locative) meaning, but in the eldest attestations of the construction, the status of the verb fluctuates between that of a semi-auxiliary (or 'quasi-auxiliary') and a full auxiliary verb, maintaining resp. a higher and lower degree of syntactic independence and relationship with the original locative meaning. See the definition by Heine (1993, p. 15): "Quasi-auxiliaries may be characterized as verbs that in most respects behave like full verbs but, when governing nonfinite (participial, gerundival, infinitival, etc.) verbs, tend to assume a grammatical or "formulaic" function." This is further illustrated through the Examples (8) and (9) below.

In order to get a general idea of the diachronic development of the inchoative construction, the analysis sets off with a general productivity measurement in Section 3.1. The following sections successively present and discuss the results of the analysis of the variables outlined in Section 2.2 above.

3.1 General productivity measure: Standardized token frequency over time

Token frequency has unequivocally been defined as the frequency of occurrence of a linguistic entity (be it a morpheme, a lexeme, a string or a construction) in running text. A high and increasing token frequency of an item would be both "a concomitant of its grammaticalization" (Thompson & Mulac, 1991, p. 319), and also a trigger for upcoming formal and functional changes of the item (Hopper & Traugott, 2003 [1993]; Bybee, 2003, p. 605). Processes of constructional changes have also been found to coincide with a rise in text frequency (Guardamagna, this volume; Rostila, 2004; Noël, 2007; Traugott & Trousdale, 2013). If the inchoative construction has developed as a schema, it should have become more frequent over time.

With these considerations in mind, Table 2 summarizes the evolution of the standardized token frequency of the inchoative construction with *poner* and *meter*, calculated per 10,000 occurrences of the verbs in the corpora.[9] Figure 2 provides a visual account of the diachronic evolution.

9. We are well aware of the fact that the standardized token frequency is mostly normalized with regard to the total corpus size, which in this case would be the sample size per time slot. However, unfortunately, CORDE and CORPES XXI do not provide sufficient information in order to calculate these sample sizes. The only information that can be retrieved is that CORDE contains 125,000,000 words in total of which 92,500,000 are representative of peninsular Spanish. Some further information is provided on the total number of words included in three time periods: (1) the Middle Ages (ca. 1200–1492): 19,425,000, (2) the Golden Age (1493–1713): 25,900,000, and (3) the contemporary period (1714–1974): 47,175,000. CORPES XXI contains 67,500,000 words for Present-day peninsular Spanish. However, given that the periodization of the corpora do not coincide with those of our samples, these data do not allow us to calculate the normalized token frequency of the inchoative construction for each time period. Therefore, an alternative way of calculating the standardized token frequency per 10,000 tokens of each verb is proposed. For instance, the 15th century contains 42,934 tokens of the verb *poner*, among which 115 are inchoative constructions. This gives a standardized frequency of 27 ((115/42,934)*10,000) in round figures. Similarly, the 19th century contains 5355 tokens of *meter*, among which 144 are inchoative constructions. The standardized frequency for *meter* in that century is 269 in round figures ((144/5355)*10,000). We believe that this method has the advantage of showing the evolution of the relative frequency of the inchoative construction compared to other uses of the verbs. Otherwise, the image of an increasing (or decreasing) use of the inchoative construction could be skewed by the general increase (or decrease) in the use of

Table 2. Standardized frequencies of tokens per 10k words

	inchoative$_{poner}$	inchoative$_{meter}$	mean
13th c.	7	85	46
15th c.	27	64	45.5
17th c.	112	115	113.5
19th c.	277	269	273
21st c.	509	64	286.5

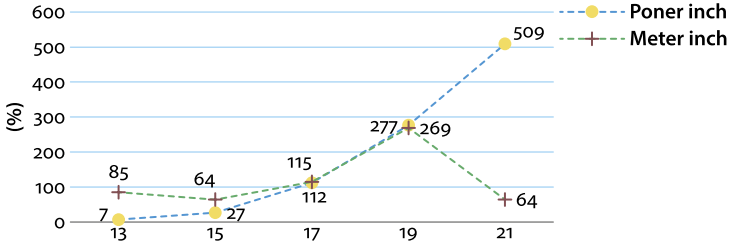


Figure 2. Evolution of the standardized token frequency

The quantitative data permit the following interesting conclusions. First, the mean standardized frequencies at the right-hand of the table show that, in general, the use of the inchoative construction has expanded over time (mostly with regard to the other uses of the verbs, but probably also in general), with important leaps between the 15th, 17th and 19th c. periods. This observation is a first (but cautious) argument in favor of the development of the inchoative construction as a schema. Second, the inchoative construction is already attested with both *poner* and *meter* in the 13th century. This suggests that the present corpus does not really show the proper 'inception' of the inchoative construction in Spanish, and that its roots must be found in earlier decades.[10]

Third, the number of occurrences with *meter* markedly outranked those of *poner* in the 13th century (85 per 10,000 tokens of *meter* in the 13th c. vs. only 7

the verbs (or one of both verbs), which is hardly informative on the evolution of the particular inchoative construction in itself. The quantitative results presented below must, however, be interpreted with regard to the other uses of the verb. But even then, we believe that an increasing relative frequency of use of the inchoative construction points towards a higher entrenchment of the construction as a schema in the speaker's mind.

10. However, neither consultation with specialists nor basic works in the field of Latin linguistics has yielded any concrete information with regard to the inception of the analytic inchoative construction in (proto-) Romance. In Latin, inchoative aspect was mainly expressed by the suffix *-scĕre* (as in *florescĕre*, 'to bloom').

for *poner*). This is in line with the etymological meaning of *meter*, which is more compatible with the goal-oriented meaning of the inchoative construction (see Section 2.2). In present-day Spanish, however, the situation is very the reverse, the relative frequency of the inchoative construction with *poner* being 8 times greater than that with *meter*. Indeed, with *poner* the inchoative construction exhibits a strong and systematic increase over time. The evolution of the construction with *meter* is somewhat different. Its frequency reaches its highest point in the 19th century, after having evolved in parallel with the construction with *poner*, but shows a remarkable drop in the 21st-century corpus.

Moreover, not only has the inchoative with *meter* become less frequent in present-day Spanish; semantically speaking, it now constitutes the marked option. Its use often has a negative connotation, and transmits the idea of a certain imprudence (3a), or of a lack of inclination (3b) or ability (3c) to perform the action expressed by the infinitive.

(3) a. ¿Por qué *se metería a husmear* en donde nada se le había perdido? [11] (CORPES XXI)
'Why would he *start to nose around* (lit. *put himself to nose*) in things that are none of his business?'

b. No ando ahora con ánimo de *meterme a hablar* a fondo del asunto […]. (CORPES XXI)
'I am not in the mood now to *start talking* (lit. to *put myself to talk*) exhaustively about the subject […].'

c. Bardem terminó COU con esfuerzo y *se metió a estudiar* Artes y Oficios porque lo que quería era ser pintor, después trabajó como portero de discoteca y bailarín de streaptease. (CORPES XXI)
'Bardem finished COU with difficulty and *started to study* (lit. *put himself to study*) Arts and Occupations because he wanted to be a painter. Afterwards he worked as bodyguard in a disco and as a striptease dancer.'

To conclude, the data suggest that, in its early use, the inchoative construction overwhelmingly preferred the verb *meter* as an auxiliary verb. In the 17th and 19th centuries, the two verbs seem to occur to a comparable extent in the construction, whereas in present-day Spanish *meter* has become a rather exceptional option (in comparison to other uses of *meter*).[12] Over time, the inchoative construction has

11. All examples cited in this article are selected from the CORDE and CORPES XXI corpora. More detailed bibliographical information for each case can easily be retrieved from these databases and is, for practical reasons, not reproduced in the text.

12. A previous study (Comer et al., 2015, p. 366) has indeed shown that in present-day Spanish, *meter* mainly occurs in the full lexical locative construction, whereas *poner* occurs more

thus become more productive (in quantitative terms), and this is mainly due to the increased use of *poner*.

3.2 Variation in the prepositional slot

A second parameter to be addressed relates to changes in structural variability due to variation in the prepositional slot. Strikingly, the CORDE-corpus attests not only cases where the INF slot is preceded by the preposition *a* ('to'), but also those preceded by the preposition *en* ('in'). This is illustrated by the allostructions in Examples (4–5).[13]

(4) a. no *se meta* **en** *escreuir* letras amorosas ni plazenteras […].
(CORDE, 15th c.)
'*Don't start to write* (lit. *put yourself* **into** *writing*) love letters or jokes.'

b. Palabras, en que muestra la poca noticia que tenia de esta materia, vicio en que pecan de ordinario los que *se meten* **á** *escribir* lo que no es de su facultad, por doctos que sean. (CORDE, 17th c.)
'Words, by which he shows how little he knew about this matter, an error usually committed by those who *start to write* (lit. who *put themselves* **to** *write*) about what they do not know about, no matter how intelligent they are.'

(5) a. por tu arte & industria *te pones* **en** *contenplar* las industrias & maneras ençima dichas. (CORDE, 15th c.)
'For your art and skills, *you start to contemplate* (lit. *you put yourself* **in** *contemplating*) the skills and the above-mentioned manners.'

b. […] *me puse* ***á*** *contemplar* semejantes maravillas.
(CORDE, 17th c.)
'[…] *I started to contemplate* (lit. *I put myself* **to** *contemplate*) similar miracles.'

The presence of the preposition *en* as a structural variant of the inchoative construction needs some additional comments. It shows, more particularly, that the inchoative construction relates to the higher level caused-motion schema through a metaphorical inheritance link (Goldberg, 1995). That is, instead of putting a

---

frequently in constructions with more grammatical meaning such as the inchoative, causative and copular constructions.

13. The corpus also documents six examples of the inchoative construction without a preposition, which have possibly been produced under the influence of the phenomenon of phonotactics: in informal, often spoken, language, the preposition is assimilated to a previous vowel ('*se ponía a pensar*' 's/he starts to think' is then pronounced as '*se ponía pensar*'). Because of their inherently different nature, these tokens have been excluded from the corpus.

participant in a concrete place, the agent or cause can 'put' a theme (oneself) in an abstract place (i.e. event, activity or a state of doing something) corresponding to the process expressed by the INF slot. It thereby illustrates the power of the metaphor CONCRETE > ABSTRACT (e.g. Traugott & Dasher, 2002; Jansegers & Enghels, 2013). As such, sentence (4a) above cautions the reader about the action of putting oneself in the process of writing love letters, and in (5a), one should 'put' oneself in the process of contemplating manners. This metaphorical extension is corroborated by examples in which an infinitive complement co-occurs with an (abstract) locative complement (*mis cosas* 'my business') (6).

(6) ¿Quién *te mete a ti en mis cosas* y *en averiguar* si soy discreto o majadero? (CORDE, 17th c.)
'Who has made you interfere in my business and decide (lit. who *puts you in my business* and *in verifying*) whether I am discrete or stupid.'

The main question now is to what extent the inchoative construction attests this structural variability both with *poner* and *meter*, and whether there is evidence of diachronic constructional change, with possibly a fixation of form. In order to answer these questions, we need to study the relative frequencies of the two structural variants for each time period. Table 3 provides the relative frequency values of the two structural variants per time period, and per auxiliary verb.

Table 3. Relative frequencies of structural variants according to the parameter [prepositions][14]

	inchoative$_{poner}$						**inchoative$_{meter}$**					
	en		**a**		**total**		**en**		**a**		**total**	
	#	%	#	%	#	%	#	%	#	%	#	#
13th c.	4	57.1	3	42.9	7	100	2	4.8	40	95.2	42	100
15th c.	12	10.5	102	88.5	114	100	5	15.2	28	84.8	33	100
17th c.	48	7.4	599	92.6	647	100	38	44.2	48	55.8	86	100
19th c.	0	–	1066	100	1066	100	21	14.6	123	85.4	144	100
21st c.	0	–	2929	100	2929	100	3	4.2	69	95.8	72	100

The quantitative results suggest that: (i) there has been a period of competition between the two structural variants with preposition *a* and *en*, both with

14. Both for *poner* and *meter*, a Fisher's Exact test confirms the statistical significance of the quantitative data, and allows us to reject the null hypothesis of independence. This implies that there is a correlation between the (independent) time variant and the (dependent) structural variation between *en* and *a*. A level of significance of 5% is accepted, which means that the correlation is considered significant when $p \leq 0.05$.

poner and *meter*, (ii) the structural variability has mainly disappeared in present-day Spanish, and (iii) the evolution of the constructions with *poner* and *meter* has been quite different.

The inchoative construction with *poner* displays a slight predominance of the preposition *en* in the 13th c. (4 out of 7 cases). From the 15th c. on, we observe a rapid increase in the amount of tokens with *a*, paralleled with a marginalization and eventual loss of the variants with *en*. As a consequence, in present-day Spanish the preposition *a* has generalized as the standard preposition with *poner*. Yet, the evolution of the prepositions in the constructions with *meter*, at the right side of Table 3, is rather different.

Even in the earliest period of the corpus, the preposition *a* is the dominant structural variant (95.2%) with *meter*. The constructions with the preposition *en* seem to gradually expand and reach a culmination point in the 17th century, at which point examples with *en* and *a* are almost equally weighted in the corpus (*en* = 47.1%; *a* = 52.9%). Thereafter, frequencies of *en* start to decrease in favor of *a*, the preposition which is now used in the most widespread structural variant, just as with *poner*. However, as opposed to what is observed with *poner*, a few cases of the $_{meter}$ inchoative construction with *en* remain in the contemporary corpus (4.2%) (7).

(7) Siempre quise tocar el piano, pero en mi casa no había. […] Un año después, como premio, *se metieron en comprar* uno, que todavía está en casa de mi madre. (CORPES XXI)
'I always wanted to play the piano, but in my house there was no piano. […] One year later, as a reward, they *decided to buy* one (lit. *they put themselves in buying* one), which is still in the house of my mother.'

Thus as far as the preposition is concerned, there was more structural variation in Medieval (13th c.–15th c.) and Classical (17th c.) Spanish. For *poner*, this syntactic variability is mostly attested in the 13th and 15th centuries, whereas with *meter* it persists until the 17th and even 19th century.

The different starting situation of the inchoative construction with *poner* and *meter* can be explained by the etymological meaning of both verbs. As the etymon of *poner* already conveys locative meaning, it more frequently combines with the lexically specified preposition *en*. The verb *meter*, on the other hand, continues the directional and goal-oriented meaning of Latin *mĭttĕre*. Given that the inchoative$_{meter}$ construction was more frequent in the 13th c. corpus than the inchoative$_{poner}$ one, it is possible that the preposition *a* also became more frequent with *poner*, due to analogy. Later on, *meter* became more strongly associated with the locative meaning than *poner*, and specialized for internally directed movements, frequently expressed by the preposition *en* (see Section 2.2).

Nevertheless, the evolution of the inchoative construction with both verbs is characterized by the gradual predominance of *a*. This appears to have led the construction to become formally more fixed. In other words, the structural variant with the preposition *a* became more entrenched, possibly because its goal-oriented meaning was most compatible with the semantics of the inchoative construction (namely, to obtain a goal in the immediate future and to start an event). As a hypothesis we can state that this has occurred through an analogical link with other micro-constructions of the inchoative subschema, which are all construed with the preposition *a* (see the examples cited above *empezar a/comenzar a bailar* 'to start to dance').[15] Put differently, it can also be argued that the micro-constructions$_{meter/poner}$ have been sanctioned by the more general inchoative schematic construction.

In sum, the inchoative construction displays decreasing structural variability over time, and seems to have developed towards a formally more fixed schema involving the preposition *a*. However, throughout its history, a tension can be observed between the impact of the semantics of the auxiliary verb (namely the locative vs. directive meaning of *poner* and *meter* respectively in the 13th c., and later on, the directional movement meaning of *meter*), and that of the construction as a whole (namely the goal-oriented meaning of the inchoative construction) on the frequency of use of structural variants with different prepositions.

3.3 Internal word order and placement of adverbials: Syntactic incorporation and unithood

A second formal parameter relates to word order within the inchoative construction, and distinguishes structural variables in which the (semi)auxiliary verb and the infinitive do, or do not, constitute a complex predicate (Fauconnier, 1983). The phenomenon of syntactic incorporation has been described within grammaticalization theories as a concomitant of the auxiliarization of a main verb. That is, when *poner* and *meter* acquire the status of auxiliary verb in the inchoative construction (as suggested in the model of Verroens (2011) for instance), they gradually become more incorporated with the infinitive complement that follows. As a consequence of this bonding, it is difficult for other lexical elements to appear between the two verbs. Note that in the inchoative construction, the intercalation exclusively concerns adverbial complements.[16]

---

15. It would be very interesting to compare the development of the inchoative construction with *put* verbs to that of other related constructions. This falls, however, beyond the scope of this article, and is left for future research.

16. This is opposed to other constructions, such as the causative one, where the intercalation may concern the main or subordinate subject, an argument of the infinitive, or a combination of those:

From a constructional viewpoint, syntactic incorporation is related to a construction's higher degree of unithood. When no lexical element is inserted between the main verb and subordinate infinitive, the action is conceptualized as a unitary whole. When, on the other hand, the two verbs are not syntactically adjacent, two relatively independent actions are conceptualized (e.g. Achard, 1998). In this light, the difference between the structural variants (a) *se pone a leer en su cuarto* and (b) *se pone en su cuarto a leer* ('he starts to read in his room') could be paraphrased as follows: in (b) the subject goes to/puts himself into his room (act 1) to start to read there (act 2), whereas in (a) the subject goes/starts to read in his room.

Both of the two structural variants cited in Section 2.2. above are found in the corpus, with the adverbial complement between the *put* verb and the infinitive (NP $_1$ + V + ADV + (PREP + INF)) (8a) or after the infinitive ((NP $_1$ + V + (PREP + INF) + ADV) (8b).

(8) a. E los moros firieron e mataron muchos christianos que *se metían* **por las casas** *a robar* . (CORDE, 15th c.)
'And the Moors injured and killed many Christians as they *started to steal from the houses* (lit. *put themselves* **in the houses** *to steal*).'

b. Cuando su huerta, que es muy mala, le produce poco, *se mete a robar* **en la del vecino** ! (CORDE, 19th c.)
'When his vegetable garden, which is very bad, does not bring in much, he *starts to steal from his neighbor's* (lit. *puts himself to steal* **in that of the neighbor**)!'

Table 4 details the relative frequency of the two variants.

Table 4. Relative frequencies of structural variants according to the parameter [+/− intercalation][17]

	inchoative$_{poner}$						**inchoative$_{meter}$**					
	+ intercal		**− intercal**		***total***		**+ intercal**		**− intercal**		***total***	
	#	**%**	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
13th c.	2	28.6	5	71.4	7	100	14	32.6	29	67.4	43	100
15th c.	14	12.2	101	87.8	115	100	5	15.2	28	84.8	33	100
17th c.	78	11.6	570	88.4	648	100	14	16.3	72	83.7	86	100
19th c.	95	8.8	971	91.2	1066	100	22	15.3	122	84.7	144	100
21st c.	99	3.4	2830	96.6	2929	100	38	52.8	34	47.2	72	100

puso *muchas vezes sus doctas manos* a scriuir el verso heliconio (CORDE, 15th c.)
'that he made often his erudite hands write the verse of Helicon'

The data clearly show that the inchoative construction with *poner* presents an uninterrupted and increasing tendency towards the selection of the variant in which the adverbial complement is placed after the unit [*ponerse* + Inf], going from 71.4% in the 13th c. to 96.6% of all cases in present-day Spanish (9a). The highly exceptional (3.4%) intercalation of elements between the unit [*ponerse* + Inf] in present-day Spanish, points towards a high degree of unithood of the construction as a whole. The inchoative with *meter* presents a rather different evolution. Up to the 15th century, an increased relative frequency of the incorporated variant is observed, whereupon it remains at a high level of over 80% until the present-day Spanish corpus, where the structural variant with intercalation has become slightly more dominant (52.8%) (9b).

(9) a. Casi *te ponés a llorar* **allí** , porque me ha visto fumándome un porro, me va a suspender. (CORPES XXI)
'You almost *start to cry there* (lit. *you put yourself to cry* **there**), because he has seen me smoking a joint, he will fire me.'

b. […] así que *me metí* **en el saco de dormir** *a planear* un suicidio sonado, y planeé dos o tres y ya me eché a llorar. (CORPES XXI)
'[…] So I got into the sleeping bag to plan (lit. *put myself* **in the sleeping bag** *to plan*) a sensational suicide, and I planned two or three of them and started to cry.'

Thus, after having developed in a parallel manner in this respect, the inchoative constructions with *poner* and *meter* have started to display a rather dissimilar behavior. This seems to go hand in hand with the development of the frequency of use of the two variants as displayed in Table 2. The construction with *poner* seems to have spread, and to have developed towards a constructional schema (including less prepositional variation and more unithood), whereas the construction with *meter* has fallen into relative disuse, possibly under the influence of the semantics of the verb. Being more semantically specialized, this verb appears with the directional *en* preposition, and more frequently maintains the semantic autonomy of both the *put* and the subordinate event.

With regard to the compositionality of the construction, a decrease of lexical intercalation is observed for the construction with *poner*, as opposed to *meter*. In Example (9a) the import of the meaning of the individual lexical items is less transparent, as the construction as a whole conveys the inchoative meaning, while

17. A Chi Square test confirms the statistic significance of the data, both for *poner* ($\chi^2 = 106{,}892$, $df = 4$, $p = 0.000$) and *meter* ($\chi^2 = 43.83$, $df = 4$, $p = 1e{-}8$). The null hypothesis according to which the time variable and the observed structural variation are independent variables can therefore be rejected.

in (9b) the locative meaning of *meterse* still filters through, and contributes to a more compositional meaning for the construction.

The analyses in this section suggest that the inchoative construction with *poner* has developed as a schema, including reduced structural variation and compositionality. The construction with *meter*, on the other hand, seems to have reversed this tendency over the last century or so. The next section examines the extent to which these findings are confirmed by the development of the semantics of the construction.

3.4 Semantic types

As specified in the definition given under (1) above, the prototypical inchoative construction consists of an animate NP_1 who gets involved in a dynamic event. Indeed, the inchoative event requires an effort, decision or intentionality on the part of the subject (10a). As a consequence, dynamic infinitives, preferably of the durative aspect, are mostly attracted to the INF slot of the construction. These are mostly activities or accomplishments, whereas states are in general ineligible. Lamiroy (1987, p. 279) calls this the "stative constraint", (10b-c), (see also Gómez Torrego, 1999; García Fernández et al., 2006; Aparicio et al., 2014).

(10) a. <u>*Se ha metido a hablar*</u> de algo que no conoce y que hace muy mal en hablar. (García Fernández et al., 2006, p. 204)
'*S/he's started to talk* (lit. <u>*put to talk*</u>) about something s/he doesn't know and which one should not talk about.'

b. * <u>*Se metían a tener*</u> hambre. (García Fernández et al., 2006, p. 203)
*'They *started to be* (lit. <u>*put themselves to be*</u>) hungry.'

c. * <u>*Se pone a ser*</u> de noche. (García Fernández et al., 2006, p. 222)
*'It *starts to be* (lit. <u>*puts to be*</u>) evening.'

It has been argued that if a constructional schema develops, speakers abstract away from individual semantic types and the schema attracts new ones. This phenomenon has been defined in terms of the productivity of the construction (see Himmelmann, 2004; Barðdal, 2008; Hilpert, 2013; Traugott & Trousdale, 2013 among others). The notion of productivity has been operationalized in very different ways in the literature, and its application to real corpus data requires advanced quantitative methods (see, for instance, Zeldes, 2012 for an overview). Given that such a proposal falls beyond the scope of the present study, we consider here only the development of the different semantic types that can be expressed by the inchoative construction. In other words, to what extent does the corpus document a semantic development of the inchoative construction which goes beyond the prototypical dynamic semantic schema?

Taking into account the animateness of the subject NP $_1$ (animate vs. inanimate vs. impersonal), and the dynamicity of the infinitive,[18] six semantic types can be distinguished in the corpus.

- type 1: NP $_1$ [ANIM] + INF [DYN]
- type 2: NP $_1$ [ANIM] + INF [-DYN]
- type 3: NP $_1$ [−ANIM] + INF [DYN]
- type 4: NP $_1$ [−ANIM] + INF [−DYN]
- type 5: NP $_1$ [IMP] + INF [DYN]
- type 6: NP $_1$ [IMP] + INF [−DYN]

Table 5 presents the relative frequencies of the six semantic types per time period for each of the two verbs.

Table 5. Relative frequencies of semantic types

	inchoative$_{poner}$													
	1		**2**		**3**		**4**		**5**		**6**		**total**	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%
13th c.	5	71.4	2	28.6	–	–	–	–	–	–	–	–	7	100
15th c.	109	94.8	3	2.6	3	2.6	–	–	–	–	–	–	115	100
17th c.	621	95.8	14	2.2	12	1.9	1	0.2	–	–	–	–	648	100
19th c.	1028	96.4	14	1.3	20	1.9	4	0.4	–	–	–	–	1066	100
21st c.	2780	94.9	18	0.6	86	2.9	26	0.9	2	0.1	17	0.6	2929	100
	inchoative$_{meter}$													
	1		**2**		**3**		**4**		**5**		**6**		**total**	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%
13th c.	43	100	–	–	–	–	–	–	–	–	–	–	43	100
15th c.	31	94	1	3	1	3	–	–	–	–	–	–	33	100
17th c.	81	94.2	5	5.8	–	–	–	–	–	–	–	–	86	100
19th c.	135	93.8	6	4.2	3	2.1	–	–	–	–	–	–	144	100
21st c.	68	94.4	3	4.2	–	–	1	1.4	–	–	–	–	72	100

18. Transitive infinitives (e.g. *write, give*) present the highest degree of dynamicity (or transfer of energy), followed by the unergative infinitives (e.g. *work, dance*), unaccusative (e.g. *fall, break*) and copulative ones (e.g. *be, become*). Therefore, transitives and unergatives have been defined as dynamic infinitives, whereas unaccusatives and copula are regarded as non-dynamic infinitives (Enghels, 2007).

Table 5 tells us that, in the 13th century, the inchoative construction with *poner* was indeed mainly restricted to the prototypical semantic type (Type 1: NP_1[ANIM] + INF[DYN]) (11a). This prototype remains the dominant one in present-day Spanish (11b) (94.4% of all cases). Diachronically, the construction with *poner* opens up, albeit to a limited extent, to other semantic types including inanimate (11c) (Types 3 and 4) and impersonal subjects (Types 5 and 6) (11d) and/or less dynamic infinitives (Types 2, 4 and 6) (11e). In the 21st-century corpus, all six semantic types occur, but the relative frequencies of the non-prototypes remain very low.

(11) a. Aristotil fue disciplo de Plato como es dicho. & *puso en aprender* […]. (CORDE, 13th c.)
'Aristotle was a pupil of Plato, as is affirmed. & he *started to learn* (lit. *put himself to learn*) […].'

b. Topo se lavó las manos y luego *se puso a barrer y fregar* . (CORPES XXI)
'Topo washed his hands and then *started to sweep and to mop* (lit. *put himself to sweep and mob*).'

c. Su labio inferior *se puso a temblar* por su cuenta. (CORPES XXI)
'His under lip *started to tremble* (lit. *put itself to tremble*).'

d. *Se puso a diluviar* . (CORPES XXI)
'It *started to rain* (lit. *put itself to rain*) heavily.'

e. Manuel continuó trabajando con más fuerza. Hasta que, con el escoplo, se arañó un dedo que *se puso a sangrar* . (CORPES XXI)
'Manuel continued working with increased force until, with the chisel, he scratched a finger that *started to bleed* (lit. *put itself to bleed*).'

The evolution of the inchoative construction with *meter* is quite similar: the prototypical Type 1 (NP_1[ANIM] + INF[DYN]) was the only semantic schema encountered in the 13th century data, and has remained the dominant one throughout the history of the construction (12a). However, as opposed to *ponerse*, almost no expansion to non-prototypical contexts can be observed: the inchoative construction does not occur with impersonal subjects (Types 5 and 6), and only a small scattering of examples with an inanimate subject (12b) (Types 3 and 4) and/or a non-dynamic infinitive (Types 2, 4 and 6) (12c) are attested.[19]

19. Given the high amount of empty cells and low frequencies, one should be cautious when applying a statistical test to these tables. A Chi Square test which excludes these cells points towards a significant correlation between the time variable and the semantic type variable for *poner* ($\chi^2 = 90{,}084$, $df = 20$, $p = 0.000$), but not for *meter* ($\chi^2 = 11{,}705$, $df = 20$, $p = > 0.05$).

(12) a. ¡Cuando su huerta, que es muy mala. le produce poco, *se mete a robar* en la del vecino! (CORDE, 19th c.)
'When his vegetable garden, which is very bad, doesn't bring in much, he *starts stealing* (lit. *puts himself to steal*) from his neighbor's!'

b. ¿Qué monstruos, qué oscuridad extraña *se ha metido a vivir* en tus noches (CORPES XXI)
'which monsters, what strange darkness *has started to live* (lit. *has put itself to live*) in your nights?'

c. La gran doña Catalina de Moncada iba primera, belleza tan peregrina que no hay alma que no quiera *meterse a ser* su esclavina. (CORDE, 17th c.)
'The great doña Catalina de Moncada went first, such an exceptional beauty that there is not a man alive who would not want to be (lit. to *put himself to be*) her shoulder-cape.'

To conclude, although the quantitative differences are rather small, the inchoative construction with *poner* seems to have opened up to less prototypical semantic types that with *meter*, suggesting its higher degree of schematicity also at the meaning side of the construction.

4. Conclusion

The above data analysis leads to a number of relevant insights situated at descriptive, theoretical and methodological levels.

To begin with, the results have empirically verified the main drawbacks, discussed in Section 1, of studying the inchoative periphrasis in Spanish solely from a grammaticalization viewpoint, and, conversely, the advantages of the constructional approach. First, by explicitly addressing the development of the formal and semantic features of the inchoative construction, it is possible to arrive at some significant generalizations showing that the inchoative periphrasis with *poner* has, to some extent, developed in a similar way to that with *meter*. Second, the historical development of the inchoative construction does not conform to a straightforward unidirectional model and can only be understood by examining the changing relationship between different structural variants over time. A constructional approach thus offers a valuable complement to the grammaticalization approach.

As a central research question, we asked whether it would be valid to claim that, over time, the inchoative construction with *put* verbs in Spanish has developed as a constructional schema. In order to answer this question, the historical evolution of alternative patterns has been examined in detail. These are the main findings.

- Over time, in terms of frequency of use, the inchoative construction has become more productive.
- As regards the structural variability defined by the prepositional slot, a general tendency has been observed towards less variability. The semantically neutral preposition *a* has become the overwhelmingly dominant option.
- The inchoative construction has gradually evolved towards more syntactic incorporation, corresponding to a higher degree of unithood of the construction as a whole.
- Although the dynamic semantic prototype remains dominant, the construction seems to have opened up (although to a limited extent) to include other semantic types.

In line with the basic tenets of diachronic constructional theories (as defined by Hilpert, 2013 and other works cited in this article), these findings suggest that a leveling process has taken place, and that an abstract constructional inchoative schema has been formed.

However, this conclusion comes with an important caveat, since it applies largely to the inchoative construction with *poner*. Although some parallelisms with *meter* could be observed in the corpus, the construction with this verb displays a rather different behavior, especially in recent times.

- With regard to its productivity, it has suffered a considerable drop in frequency since the 19th c.
- Although the variable using the preposition *a* has become the dominant one over time, examples can be found with the locative *en*.
- In present-day Spanish, the structural variables displaying lack of syntactic incorporation and unithood have become as common as those which display it.
- The construction with this verb has not opened up to include semantic types other than the prototypical to any significant degree.

In sum, despite similarities between the inchoatives with *poner* and *meter* at the constructional level, this study shows that the impact of the core lexical items on the evolution of a construction cannot be understimated. Only the inchoative periphrasis with *poner* seems to have developed as a constructional schema. Given that *poner* is the semantically most neutral member of the quasi-synonymous verb pair, the inchoative construction it forms seems to have invaded the domain of that formed with the semantically specialized verb *meter*.

The conclusion that Spanish has developed a schema for expressing inchoativity with *put* verbs should therefore be qualified: at the constructional level, there is a tendency towards reduced variation and compositionality, but a semasiolog-

ical perspective which examines the impact of individual lexical items offers an additional and more nuanced account. To conclude, the case study described in this article suggests that grammaticalization and diachronic constructional studies should go hand in hand.

References

Achard, M. (1998). *Representation of cognitive structures. Syntax and semantics of French sentential complements*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110805956

Aparicio, J., Coll-Florit, M., & Castellón, I. (2014). Perífrasis incoativas: aproximación cognitiva y estudio de corpus. *Sintagma*, 26, 73–88.

Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8

Bergs, A., & Diewald, G. (Eds.). (2008). *Constructions and language change*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211757

Bybee, J. (2003). Mechanisms of change in grammaticization: the role of frequency. In J. Brian, & R. Janda (Eds.), *The Handbook of Historical Linguistics* (pp. 602–623). Oxford: Blackwell. https://doi.org/10.1002/9780470756393.ch19

Bybee, J., & Torres Cacoullos, R. (2009). The role of prefabs in grammaticization. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatly (Eds.), *Formulaic Language* (pp. 187–217). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.82.09the

Cano Aguilar, R. (1992). *El español a través de los tiempos*. Madrid: Arco/Libros.

Cappelle, B. (2006). Particle Placement and the Case for 'Allostructions'. *Constructions*, 1, 1–28.

Cifuentes Honrubia, J. L. (1999). *Sintaxis y semántica del movimiento*. Alicante: Instituto de Cultura Juan Gil-Albert.

Comer, M., Enghels, R., & Vanderschueren, C. (2015). De verbos de colocación a pseudo-copulativos: procesos de gramaticalización en los cuasi-sinónimos *poner/meter* y *pôr/meter* en español y en portugués. *Zeitschrift für romanische Philologie*, 131(2), 355–382. https://doi.org/10.1515/zrp-2015-0023

Comer, M., & Enghels, R. (in press). La evolución de las perífrasis verbales causativa e incoativa con *poner* en español: cambios de construcción y extensiones metafóricas. *Bulletin of Hispanic Studies*.

Corominas, J., & Pascual, J. A. (Eds.). (1980–1991). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.

Coussé, E. (2018). Grammaticalization, host-class expansion and category change. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category change from a constructional perspective*. Amsterdam: John Benjamins.

Croft, W. (2001). *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Enghels, R. (2007). *Les Modalités de perception visuelle et auditive: différences conceptuelles et répercussions sémantico–syntaxiques en espagnol et en français*. Tübingen: Niemeyer. https://doi.org/10.1515/9783110944884

Fauconnier, G. (1983). Generalized Union. In L. Tasmowski, & D. Willems (Eds.), *Problems in syntax* (pp. 195–229). Gent: Plenum. https://doi.org/10.1007/978-1-4613-2727-1_9

Garachana Camarero, M. (2016). Restricciones léxicas en la gramaticalización de las perífrasis verbales. *RILCE*, 32(1), 134–158.

García Fernández, L. (Eds.). (2006). *Diccionario de perífrasis verbales*. Madrid: Gredos.

Goldberg, A. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gómez Manzano, P. (1992). *Perífrasis verbales con infinitivo (valores y usos en la lengua hablada)*. Madrid: Universidad Nacional de Educación a Distancia.

Gómez Torrego, L. (1999). Los verbos auxiliares. Las perífrasis verbales de infinitivo. In I. Bosque, & V. Demonte (Eds.), *Gramática Descriptiva de la Lengua Española* (pp. 3323–3441). Madrid: Espasa.

Guardamagna, C. (this volume). Type frequency, productivity and schematicity in the evolution of the Latin *secundum* NP construction.

Heine, B. 1993. *Auxiliaries, Cognitive Forces, and Grammaticalization*. New York: Oxford University Press.

Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 83–102). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Hilpert, M. (2008). *Germanic Future Constructions. A usage-based approach to language change*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.7

Hilpert, M. (2013). *Constructional Change in English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139004206

Hilpert, M. (this volume). Three open questions in Diachronic Construction Grammar.

Himmelmann, N. (2004). Lexicalization and grammaticalization: opposite or orthogonal? In W. Bisang, N. Himmelmann, & B. Wiemer (Eds.), *What Makes Grammaticalization – A Look from its Fringes and Its Components* (pp. 19–40). Berlin: Mouton de Gruyter.

Hopper, P., & Traugott, E. (2003 [1993]). *Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Ibarretxe, I. (2012). Placement and removal events in Basque and Spanish. In A. Kopecka, & B. Narasimhan (Eds.), *Events of putting and taking: A crosslinguistic perspective* (pp. 123–143). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.100.10ant

Jansegers, M. (2015). *Hacia un enfoque múltiple de la polisemia. Un estudio empírico del verbo multimodal 'sentir' desde una perspectiva sincrónica y diacrónica*. Dissertation. Gent: Universiteit Gent.

Jansegers, M., & Enghels, R. (2013). De verbo de percepción a marcador de disculpa: la evolución diacrónica del verbo sentir en español. *Revue de Linguistique Romane*, 77(305), 139–166.

Lamiroy, B. (1987). The complementation of aspectual verbs in French. *Language*, 63(2), 278–298. https://doi.org/10.2307/415657

Lamiroy, B., & De Mulder, W. (2011). Degrees of Grammaticalization across Languages. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 302–317). Oxford: Oxford University Press.

Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of language*, 14(2), 177–202. https://doi.org/10.1075/fol.14.2.04noe

Real Academia Española: Banco de datos (CORDE). *Corpus diacrónico del español*. http://corpus.rae.es/cordenet.html

Real Academia Española: Banco de datos (CORPES XXI). *Corpus del Español del Siglo XXI (CORPES)*. http://web.frl.es/CORPES/view/inicioExterno.view

Rostila, J. (2004). Lexicalization as a way to grammaticalization. In F. Karlsson (Ed.), *Proceedings of the 20th Scandinavian Conference of Linguistics (Helsinki)*. http://www.ling.helsinki.fi/kielitiede/20scl/Rostila.pdf.

Thompson, S. A., & Mulac, A. (1991). The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of pragmatics*, 15, 237–251. https://doi.org/10.1016/0378-2166(91)90012-M

Torres Cacoullos, R. (2000). *Grammaticization, Synchronic Variation, and Language Contact: A Study of Spanish Progressive '-ndo' Constructions*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.52

Traugott, E., & Dasher, R. (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, E., & Trousdale, G. (2013). *Constructionalization and constructional change*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Trousdale, G. (2008). Constructions in grammaticalization and lexicalization. Evidence from the history of a composite predicate construction in English. In G. Trousdale, & N. Gisborne (Eds.), *Constructional approaches to English grammar* (pp. 33–67). Berlin: Mouton de Gruyter.

Trousdale, G. (2014). On the relationship between grammaticalization and constructionalization. *Folia Linguistica*, 48(2), 557–577. https://doi.org/10.1515/flin.2014.018

Vergara Wilson, D. (2009). From 'remaining' to 'becoming' in Spanish: The role of prefabs in the development of the construction 'quedar(se) + adjective. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatly (Eds.), *Formulaic Language* (pp. 273–295). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.82.13fro

Verroens, F. (2011). *La construction inchoative se mettre à: syntaxe, sémantique et grammaticalisation*. Dissertation. Gent: Universiteit Gent.

Zeldes, A. (2012). *Productivity in Argument Selection. From Morphology to Syntax*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110303919

Noun constructions

Reduction or expansion? *A bit of* both

A case study on the development of German degree modifiers

Jakob Neels & Stefan Hartmann
Leipzig University | University of Bamberg

This paper presents a diachronic analysis of the German quantifier/degree-modifier constructions *ein bisschen* ('a bit$_{DIM}$') and *ein wenig* ('a little'). On the basis of data from two historical corpora, we examine to what extent these constructions followed a grammaticalization path comparable to the one Traugott (e.g. 2008a) observed for their English counterparts. Quantitative analyses on their gradual context expansion are combined with qualitative analyses on potential bridging contexts for reanalysis. Their combined histories suggest that the older German modifiers may have served as attractor sets for an increasingly strengthened mid-level quantifier/degree-modifier schema, promoting the emergence of younger modifiers. These observations stress the crucial role that network links between constructions can play in grammaticalization.

Keywords: degree modifiers, analogy, paradigmatization, low-frequency grammaticalization, German

1. Introduction

Both in grammaticalization studies and in diachronic construction grammar, the development of degree modifiers such as *a bit* or *a little* have sparked much interest (e.g. Traugott, 2008a, 2008b; Brems, 2011; Traugott & Trousdale, 2013). It is not hard to see why, in this particular case, both perspectives complement each other. On the one hand, the development of quantifiers from a semantically fairly specific noun like *bite* (e.g. *of bread, of food*) to a partitive (*a bit of the food*), quantifier (*a bit of money*) and degree modifier (*a bit shy*) is a prime example of key grammaticalization processes such as semantic bleaching (Hopper & Traugott, 2003; Lehmann, 2015) or context expansion (Himmelmann, 2004). On the other hand, patterns like [*a bit of* N], [*a shred of* N], or the more abstract [NP *of* NP] are

https://doi.org/10.1075/cal.21.c6

textbook examples of constructions, i.e. form–meaning pairs at various levels of abstraction (cf., e.g., Croft, 2001; Goldberg, 2006). What is more, these constructions lend themselves particularly well to studying the complex relations between constructions, thus giving a glimpse into the organization of the so-called constructicon (cf., e.g., Hilpert, 2014). In Goldberg's (1995) terms, inheritance relations can be assumed between the more abstract [NP *of* NP] constructions and the subschemas in which the first NP slot is filled by a combination of an indefinite determiner and one specific noun.

So far, many studies on the development of quantifier/degree-modifier constructions have focused on English (e.g. Traugott, 2008a, 2008b; Brems, 2007, 2012; De Clerck & Brems, 2016). However, other languages also have quantifier/degree-modifier constructions that are strikingly similar to the English ones, as is the case, for instance, in Dutch (e.g. De Clerck & Colleman, 2013) and in Spanish (e.g. Verveckken, 2012, 2015). Therefore, it seems worthwhile investigating whether these have undergone a similar development as their English counterparts. In this paper, we will focus on the German constructions *ein wenig* 'a little' and particularly *ein bisschen* 'a bit$_{DIM}$'[1] and their development at the beginning of the New High German period, which, according to the most widespread periodization scheme of the German language history, begins in the mid-17th century. While *ein wenig* had already been well established by then, (*ein*) *bisschen* was a fairly new construction. It is first attested in the 15th century (cf. Pfeifer, 1993) and still rather infrequent at the beginning of the time period under investigation. Importantly, the paradigmatic relations that can be assumed to hold between both variants also extend to other, less frequent constructions such as *ein Quäntchen* 'a quantum$_{DIM}$ (of)', *eine Handvoll* 'a handful (of)' or the much older variant *ein lützel* 'a little', which, however, has fallen out of use in most variants of German long before the beginning of the New High German period. While many of the previous studies on similar binominal constructions in Spanish, Dutch and English focus on their development into grammaticalized quantifiers, the present case study is equally concerned with a further step of grammaticalization, namely the change from quantifier into degree modifier.

This study addresses three questions with regard to the diachronic development of the most frequent German quantifier/degree-modifier constructions *ein wenig* and (*ein*) *bisschen*:

1. Subscript DIM stands for 'diminutive', as marked by the German suffixes *-chen* and *-lein*, for instance. Other subscript glosses used in this paper are FEM for 'feminine' and NEUT for 'neuter'.

a. To what extent did the German quantifier/degree-modifier constructions follow a grammaticalization path comparable to the one Traugott (e.g. 2008a) observed for their English counterparts?
b. Can the attested developments be usefully conceived of as constructional change involving both expansion and reduction?
c. How do older and younger German quantifier/degree-modifier constructions interact in a 'layered' (Hopper, 1991) domain of grammar?

To answer these questions, we will first present a brief research overview before turning to our own diachronic corpus study, which combines quantitative analyses on the gradual context expansion of *ein wenig* and *ein bisschen* with qualitative analyses on potential bridging contexts (cf. Heine, 2002). Section 2 sketches out how diachronic construction grammar brings together two contrasting research traditions that can be labelled "grammaticalization as reduction" and "grammaticalization as expansion", following Traugott & Trousdale (2013). After a review of Traugott's work on English degree modifiers in Section 3, Section 4 introduces the quantifier/degree-modifier constructions investigated in our follow-up study on German. Section 5 then describes the corpus data and methods used. The results of the corpus analyses are presented and discussed in Sections 6 to 8. Finally, Section 9 summarizes our findings and provides an outlook.

Setting up this investigation in a bottom-up fashion on the basis of a large sample of authentic usage data, we adopt a usage-based version of construction grammar. That is, we subscribe to the central tenets of usage-based theory (e.g. Bybee, 2010; Tomasello, 2003; Langacker, 2000; Hopper, 1987; Givón, 1979; see Barlow & Kemmer, 2000 for a summary): language structure is taken to be grounded in and built up from concrete usage events with the aid of domain-general cognitive abilities such as categorization, pattern extraction, analogy and inference-making; there is a constant feedback loop between grammar and usage, which holds for both the communal grammar and the mental grammar of individual speakers throughout their lifetime. This feedback loop allows grammatical changes to proceed in a gradual but directed manner. It follows from these assumptions that the organisation and reorganisation of grammar must be strongly affected by frequency of use (e.g. Diessel, 2007; Diessel & Hilpert, 2016). We will therefore pay close attention to the frequency profiles of the German quantifier/degree-modifier constructions in our account of their grammaticalization.

2. Reduction and expansion in a construction grammar framework for grammaticalization

According to Traugott (2010) and Traugott & Trousdale (2013, pp. 96–112), two contrasting approaches have developed over the past thirty years of grammaticalization research. On the one hand, there is a line of research maintaining that grammaticalization is essentially a process of reduction (e.g. Lehmann, 2015; Givón, 1979; Heine et al., 1991). In this approach, grammaticalizing expressions are said to reduce and freeze in form. They become more dependent on, and later even obligatory in particular syntagmatic contexts. Meaning changes are thought of in terms of semantic bleaching, i.e. the loss of concrete meaning. On the other hand, there is a tradition of research emphasizing that grammaticalization involves expansion. This view is perhaps most clearly represented in Himmelmann's (2004) notion of context expansion. However, already in the 1980s, some researchers stressed the expansive nature of grammaticalization, focusing on the role of pragmatic enrichment/strengthening and on the emergence of polysemy and polyfunctionality (e.g. Traugott, 1988; Sweetser, 1988; Bybee et al., 1994; Tabor & Traugott, 1998).

Construction grammar has proved a fruitful framework for studying grammaticalization and related phenomena such as lexicalization and degrammaticalization. Recent construction grammar approaches explicitly address the issue of grammaticalization-as-reduction versus grammaticalization-as-expansion. As Traugott & Trousdale (2013, p. 147) put it, "grammatical constructionalization involves expansion in construction types and range of use on the one hand, and chunking and fixing of form on the other." Reduction and fixing in form and expansion and generalization in function may be considered two sides of the same coin when conceiving of grammaticalization as a routinization process that turns more controlled into more automated linguistic activities (cf. Lehmann, 2016; Bybee, 2003a, 2010; Haspelmath, 1999, pp. 1054f.). By spelling out how reduction and expansion are intertwined, the framework of usage-based diachronic construction grammar manages to bring together the two longstanding traditions in grammaticalization research.

Moreover, construction grammar approaches to grammaticalization put emphasis on constructional networks as well as constructional levels. A new node in the taxonomic network ('constructicon') is formed, as a new more procedural construction emerges on the basis of extant, more referential/contentful constructions (cf. Goldberg, 1995; Langacker, 2008; Bybee, 2010; Traugott, 2014; Diessel, 2015). Interrelated constructions may be entrenched at various degrees of abstraction and schematicity. This idea is, for instance, captured in the concepts

of micro-, meso- and macro-constructions introduced by Traugott (2007, 2008a, 2008b):

(1) a. Macro-constructions: high-level schemas, the highest level relevant for the discussion at hand.
b. Meso-constructions: sets of similar-behaving constructions (with a shared structure).
c. Mirco-constructions: individual construction types (not instances of use).
d. Constructs: empirically attested tokens.

For illustration, consider the English constructions [*will* V], [*be going to* V], [*be about to* V], [*shall* V] and (non-standard) [*be fixing to* V] (Trousdale & Norde, 2013, p. 36). Each of these expressions constitutes an individual micro-construction, but they can also be subsumed under one macro-construction on grounds of their shared function of marking the future. Some of these future-marker micro-constructions are formally more alike than others: [*be going to* V] and [*be fixing to* V] can be said to be sanctioned by one mid-level schema; [*will* V] and [*shall* V], on the other hand, belong to a different mid-level schema. This is the level of meso-constructions in-between specific micro-constructions and highly abstract macro-constructions. Such more or less schematic units are assumed to serve as the levels at which change-enabling processes and mechanisms like analogy operate. A single autonomous micro-construction with a high discourse frequency is possibly only weakly linked to a higher-level schema (meso- or macro-constructions). However, the more members a higher-level schema has, the more entrenched and productive it typically is and the stronger the links across constructional levels tend to be. Network links between formally or functionally related constructions provide micro-constructions with possible directions of change.

3. English quantifier/degree-modifier constructions

Traugott (2007, 2008a, 2008b) demonstrates the relevance of constructional levels in her research on English degree modifiers. She studies a set of expressions with a shared binominal [NP *of* NP] structure, such as *a bit of*, *a lot of*, *a kind of*, *a sort of*, *a bunch of* and *a shred of*. For the diachronic development of these expressions, Traugott proposes the following grammaticalization path:

(2) pre-partitive > partitive > quantifier > degree modifier > free adverb

This development is assumed to have taken place mainly between the 15th and 19th centuries. Each stage is illustrated by one example of the *a bit of* construction in (3) below (adopted from Traugott, 2008b, p. 29).

(3) a. *Hu he wrec in adam* ***þe bite of*** *an eappel*
'How he avenged in Adam the bite of an apple' (c.1230, MED)
b. *Gif God was made of* ***bits of*** *breid* (c.1550, OED)
c. *if you think to scape with sending mee* ***such bitts of*** *letters you are mistaken* (1653, PCEEC)
d. *I would not be* ***a bit*** *wiser,* ***a bit*** *richer,* ***a bit*** *taller,* ***a bit*** *shorter, than I am at this Instant* (1723, Richard Steel, *The Conscious Lovers*, Internet Archive)
e. *Lucilia: Hear me. Cleontes: Not* ***a bit****.* (1739, Henry Baker, *The Cit Turn'd Gentleman*, Internet Archive)

Traugott (2008b, p. 23) argues that innovative uses with new pragmatics and semantics led to mismatches between form and meaning. This tension was resolved via syntactic reanalysis:

(4) $[NP_1 [of\ NP_2]] > [[NP_1\ of]\ NP_2]]$

In the English binominal constructions, the first NP with *bit*, *lot*, *kind*, *sort* etc. ceased to be the syntactic head and developed into a modifier (cf. *kinda*, *sorta*). An important factor in the expansion of these constructions to degree-modifier functions was the existence of quantifying degree modifiers such as *quite* and *all*, which can combine with nouns as well as with adjectives, verbs and other word classes (e.g. *quite a hypocrite* versus *quite hypocritical*). Speakers and hearers, Traugott (2007, p. 536) hypothesizes, matched "the binominal micro-construction with the structural properties of the adverbial degree modifier macro-construction, i.e., reanalysis as an adverbial, licensing syntactic contexts typical of this new function (e.g., adjectives, as in *sort of cold*, *a lot wiser*)." In short, grammaticalization progressed as language users matched and realigned construction types based on functional and structural similarity.

4. German quantifier/degree-modifier constructions: An overview

As pointed out in the Introduction, there are a number of quantifier/degree-modifier constructions in German that are strikingly similar both to each other and to the English constructions studied by Traugott. The most frequent forms *ein wenig* and *ein bisschen*, exemplified in (5) and (6), can be used quasi-synonymously in present-day German. However, their distribution seems to be partly conditioned by regional factors (cf. Tiefenbach, 1987, p. 7). In addition, the younger variant

ein bisschen was stigmatized as colloquial until the mid-19th century at least (cf. Tiefenbach, 1987, p. 6, footnote 6).

(5) *aber seine Aussprache war **ein wenig** bäuerisch, und sein Auge blickte nicht ein*
'but his pronunciation was a little rural, and his eye didn't look fine'
(1805, DeReKo-HIST)

(6) *Warte doch, du mußt noch **ein bißchen** Eigenlob hören.*
'Wait, you have to hear a bit of self-praise.' (1896, DeReKo-HIST)

Apart from these two, *ein lützel* 'a little', exemplified in (7), used to be quite frequent up to the Middle High German period but, in most varieties of German, fell out of use in the 17th century (Pfeifer, 1993).

(7) *und alse er von dem valle **ein lützel** sich erholte*
'and when he recovered a bit from the fall'
(c.1210, Gottfried von Straßburg, *Tristan*, MWB)

Lützel is also attested as a free adjective, but both in its free use and in the quantifier/degree-modifier construction it was superseded by *wenig*, which is first attested in the 9th century as an adjective meaning 'miserable, minor, little'. However, the more concrete meaning variants – such as *mit einer wênegen rindes hût* 'with a small cow hide' (12th century, Eneide, MHDBDB) – were lost over time. As *wenig* became quasi-synonymous with *lützel*, it seems reasonable to assume that the quantifier/degree-modifier construction [*ein lützel* X] served as a model for its younger counterpart [*ein wenig* X].

While *bisschen* is attested from the 15th century onwards, its 'generalized' quantifier/degree modifier use, according to Kluge (2012), only dates back to the 17th/18th century. However, a fairly similar use of the non-diminutivized form *ein bissen* can be found in negated contexts as early as in the Middle High German period.

(8) *Daz doch ist nirgen **bissen** war.*
'This is not a bit true.' (14th century, Minneburg, MHDBDB)

(9) *du tust im nyrgen **pissen** we.*
'You don't hurt him a bit' (14th century, Minneburg, MHDBDB)

In these examples, *bissen* is used to intensify the negation.[2] While it seems unlikely that such uses might have influenced the quantifier/degree-modifier use of *biss-*

2. Given the sparse data available for the Middle High German period, it cannot be determined if *bissen* used to be a pure negative polarity item. However, the use of *bissen* as a negation intensifier is interesting in view of Brems' (2007) discussion of polarity-sensitive small-size nouns such as *a shred of* and *a jot of*.

chen, which only started much later, these attestations show that the metonymic relation between *bissen* 'bite' and smallness had already been well established when the quantifier/degree-modifier use of *ein bisschen* emerged.

This paper will focus on *ein bisschen*, *ein wenig*, and the relation between both at the beginning of the New High German period. However, it should be kept in mind that these are only the two most frequent quantifier/degree-modifier constructions during that time period. Other variants exist as well, although most of them are subject to more rigorous constraints than *ein bisschen* and *ein wenig.* For instance, *eine Handvoll*, as in (10), can only be used as a quantifier with a concrete *modificatum*, not with abstract nouns and not as a degree modifier (**eine Handvoll Armut* 'a handful of poverty', **eine Handvoll schlecht* 'a handful bad'). *Ein Funke/Fünkchen/Fünklein* 'a spark (of)', by contrast, can only be used if the modified item is abstract, like *Liebe* 'love' in (11).

(10) *er war geflohen vor* ***einer Handvoll*** *waghalsiger Tyroler*
'He had fled a handful of reckless Tyroleses'
(1835, Bettina v. Arnim, Goethes Briefwechsel, DTA)

(11) *Schließt eurem Hertzen ein / wie ich /* ***ein Fünncklein*** *Liebe*
'Take into your heart, like my, a $\text{spark}_{\text{DIM}}$ of love'
(1704, Abschatz, Gedichte, DTA)

Unlike these two constructions, *ein wenig* and *ein bisschen* can be used in a broad variety of contexts, and they can function as partitives (*ein bissel Brots* 'a bite_{DIM} of $\text{bread}_{\text{GEN}}$', 1615, Albertus, Landstörtzer, DTA), quantifiers (*er bat um ein bißchen Essen* 'he asked for a bit_{DIM} of food', 1819, GRI, DeReKo-HIST), and degree modifiers (*ein bißchen ungestüm* 'a bit_{DIM} vehement', 1819, HK3, DeReKo-HIST). As the quantifier and degree-modifier uses are the most frequent ones in the corpus data and also the most prominent ones in present-day German, we refer to these constructions as quantifier/degree-modifier constructions; note, however, that the partitive reading is still available not only for *ein bisschen* but also for *ein wenig*, as in *ein wenig von dem Pulver* 'a little of the powder' (1690, Mattioli, Theatrvm Botanicvm, DTA).

If the development of German quantifier/degree-modifier constructions follows a similar path as the diachronic evolution of their English counterparts sketched in Section 3, it can be assumed that the use of *ein bisschen* and *ein wenig* as quantifiers and degree modifiers is a more recent development than their use as partitives (and, in the case of *ein bisschen*, its literal use in the sense of 'bite'). While the categorization of a specific usage variant as partitive, quantifier, or degree modifier is not always possible in a clear and straightforward way, a variety of parameters can be operationalized to assess major tendencies in the diachronic

development of the constructions under discussion. These parameters, along with our data sources, will be discussed in the next two sections.

5. Studying the German degree modifiers: Corpus data and methods

In order to investigate the diachronic development of German degree modifiers from the 17th to the 19th century, we conducted an exhaustive search in two corpora: the German Text Archive (*Deutsches Textarchiv*, DTA) and the historical archive of the German Reference Corpus (*Deutsches Referenzkorpus*, DeReKo-HIST). The DTA (www.deutschestextarchiv.de) is a collection of German texts covering the period from 1600 to 1900. Still under development, it currently consists of ca. 138 million tokens from more than 2,000 texts. The DeReKo-HIST archive is made up of ten different subcorpora comprising historical newspaper texts as well as literary works and the entire GerManC corpus (Durrell et al., 2007). Overall, it contains about 68 million tokens from the second half of the 17th century until 1962 which are, however, distributed quite unequally over the different subcorpora. Unfortunately, the corpus documentation does not feature any information about token frequencies per individual time period. In the case of the DTA, these frequencies can be easily retrieved as all corpus documents are publicly accessible in various formats. Figure 1 visualizes the distribution of the DTA texts over the different decades covered by the corpus and over the four text types the texts are assigned to. When taking a closer look at the design of the DTA, it becomes obvious that it is a rather 'opportunistic' corpus as well, as it is not balanced for different time periods and as some text types are overrepresented, while others – particularly newspaper texts – hardly occur at all in the data. However, given their sheer size, these corpora are currently the best resource available for investigating the time period in question, and especially for investigating mid- and low-frequency phenomena. As we will see below, (*ein*) *bisschen* is rather infrequent, despite its highly frequent use in (colloquial) present-day German.

From both databases, all instances of *ein wenig* and (*ein*) *bisschen* were extracted. In the case of (*ein*) *bisschen*, regional variants with other diminutive suffixes were also taken into account (e.g. *bissel*, *bissle*). Note that in the case of the older construction, only instances in which *wenig* is preceded by the indefinite determiner *ein* were searched. Just as in the case of its English counterpart *little*, *wenig* is used as an adjective outside of the [*ein wenig* X] construction, which can be used attributively or predicatively: *wenig Geld* 'little money', *sie isst wenig* 'she eats little'. Therefore, it is not possible to leave out the determiner in the quantifier/degree-modifier construction. For instance, *sie gab mir ein wenig Geld* 'she

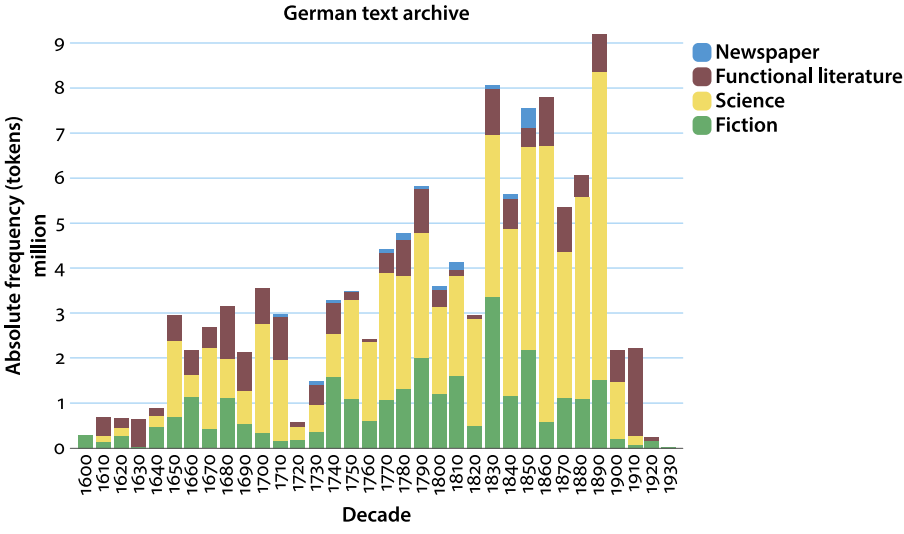


Figure 1. Distribution of the DTA texts over the individual decades and text types

gave me a little [bit of] money' cannot be replaced by *sie gab mir wenig Geld* 'she gave me little money'. The latter entails a quantitative evaluation of the amount of money taking a certain context-dependent 'norm' as reference point, expressing that the sum is considerably below this reference point. In the case of *ein wenig Geld* 'a little money', by contrast, one could say that the lower boundary (i.e. zero) is taken as the reference point. The difference is even more obvious in cases where *ein wenig* is used in a temporal sense. *Ich gehe wenig spazieren*, lit. 'I little go for a walk', means that I (habitually) hardly ever go for a walk, while *Ich gehe ein wenig spazieren*, lit. 'I go for a walk a little', means that I go for a walk for a short time. If anything, the determiner can be shortened to *'n*:

(12) *wir wollen uns den August da drüben jetzt mal **'n wenig** von nahe besehn*
'Let's take a little closer look at this guy over there'
(1910, Hauptmann, Der Narr, DWDS)

Also, *wenig* in the quantifier/degree-modifier sense cannot occur with a determiner other than *ein*. **Das wenig Geld*, for instance, is not possible; *das wenige Geld* 'the little money', however, is possible, but here *wenig* is used as a free adjective again (with a certain sum of money as reference point, as discussed above). By contrast, *bisschen* can occur with different determiners, as in (13), or even without a determiner, as in (14).

(13) *Ach, **das bischen** Leben, wieviel Angst und Qual hat man davon.*
'Oh, that bit of life, how much anxiety and pain you have from it!'
(1891, Frapan, Bittersüß, DTA)

(14) *"Ich lese **bißchen**!"*
'I read (a) bit!' (1892, DeReKo-HIST)

Therefore, all instances of *bisschen* were taken into account regardless of the preceding context, whereas for *wenig* only cases in which it occurs with the determiner *ein* were considered.

After extracting all attestations of the two patterns from the corpora, false hits – such as *ein wenig bekannter Autor* 'a little known author' – were manually discarded. Overall, 3,226 instances of (*ein*) *bisschen* and 15,783 instances of *ein wenig* were found in the data. Table 1 summarizes the token frequencies. The differences between the sum total of instances in DeReKo-HIST and DTA, on the one hand, and the joint data from DeReko-HIST/DTA, on the other, are due to the fact that some texts occur in both databases. Therefore, duplicates were removed from the joint dataset semi-manually by calculating the normalized Levenshtein distance (which indicates the number of edits necessary to change a given string *A* into string *B*) for each pair of attestations in the concordance and manually checking the rows with the lowest values. This allowed for identifying cases in which two different editions of the same text adopting slightly different graphemic conventions had entered the data. If the dates of two equivalent attestations differed, the older attestation was retained.

Table 1. Token frequencies of (*ein*) *bisschen* and *ein wenig* in the German Text Archive, the German Reference Corpus, and the joint DTA/DeReKo-HIST database

	DTA	DeReKo-HIST	DTA/DeReKo joint database
(ein) bisschen	922	2,628	3,226
ein wenig	9,176	7,382	15,783

6. Quantitative corpus analyses

This section presents the results of the quantitative analyses of the about 19,000 attestations of *ein wenig* and *ein bisschen* in the two databases. It is structured around the following three topics: token frequency, parts of speech modified by the constructions, and determiner variation.

6.1 Token frequency

Since the total number of tokens per decade in the database as a whole can only be reliably determined for the DTA, the comparison of token frequencies reported in this section relies exclusively on the DTA data, i.e. all 922 instances

of (*ein*) *bisschen* and all 9,176 instances of *ein wenig* to be found in the DTA. Figure 2 plots the normalized token frequencies of *ein wenig* and *ein bisschen*, including both non-logarithmized and log-transformed frequencies. Logarithmic transformation, as Baayen (2008, p. 39) points out, can substantially reduce potential skews in the data and thus reduce the influence of outliers. As the left figure shows, the frequency of *ein wenig* drops notably, while its younger competitor (*ein*) *bisschen* becomes more frequent over time.

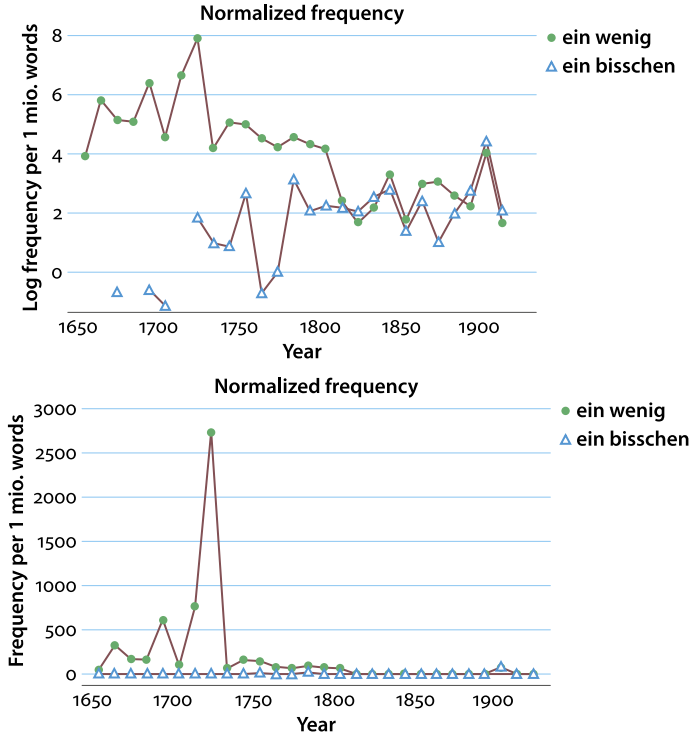


Figure 2. The left panel plots the normalized log token frequencies of *ein wenig* and (*ein*) *bisschen* in the German Text Archive (DTA). The right panel shows the non-logarithmized frequencies; here, the overall picture of the diachronic development is distorted by one extreme outlier

In addition to their descriptive value, frequency profiles of constructions can have explanatory value when interpreted with usage-based theory: Bybee (2003a, 2003b, 2010), Krug (2000, 2003), Haiman (1994) and others have put forward the hypothesis that increasing frequency of use is not only an outcome but also a motor in grammaticalization because regular repetition fuels domain-general cognitive processes, such as chunking, habituation and neuromotor automation,

that may be responsible for the formal and functional changes characteristic of grammaticalization.

6.2 Parts of speech modified by the constructions

All corpus hits were annotated for the part of speech of the *modificatum*. Both constructions can modify different word classes, which in turn might give clues to their degree of grammaticalization (viz. their status as partitive, quantifier, degree modifier, or free adverb).

For the older construction *ein wenig*, no significant change is detectable in the joint DTA/DeReKo-HIST database from about ca. 1725 onwards.[3] Before that, some ups and downs can be observed in the case of nouns and verbs, as plotted in Figure 3. However, this may be largely due to the low absolute frequencies in the period in question (see Figure 5 below). Also note that the first four decades (the first nine decades in the case of *ein bisschen* in Figure 4 below) have been discarded from the present analysis due to the very low total number of hits. After 1725, the word class modified most frequently by *ein wenig* is verbs throughout the entire period in question, with a very slight initial upward trend. For adjectives and nouns, an equally insignificant downward trend can be observed in the first few decades after 1725, but altogether, they consistently make up for about 30% and 10% of the items modified by *ein wenig*, respectively.

In the case of (*ein*) *bisschen*, a fairly clear pattern of diachronic change can be detected from about 1750 onwards (as pointed out above, the absolute numbers in the first few decades are too low to allow for a meaningful analysis of relative frequencies). While the relative frequencies of nouns in the [(DET) *bisschen* X] construction constantly decreases, adjectives and especially verbs become more frequent. Interestingly, with this diachronic shift, *ein bisschen* approximates the usage patterns of *ein wenig* with respect to the proportions of verb, adjective and noun modification.

Given the differences in absolute token frequency between the individual decades, it is useful to take a look not only at the relative values but also at the absolute frequencies of the different parts of speech. Figure 5 plots the absolute frequencies of items belonging to the three major parts of speech modified by *ein wenig* and (*ein*) *bisschen*, respectively.

The observation that nouns become less frequent in the [(DET) *bisschen* X] construction is very much in line with the grammaticalization scenario leading

3. For testing the significance of frequency changes, the signed-rank correlation test Kendall's tau was used throughout the paper (cf. Hilpert & Gries, 2009 on the use of Kendall's tau for assessing frequency changes in diachronic corpora).

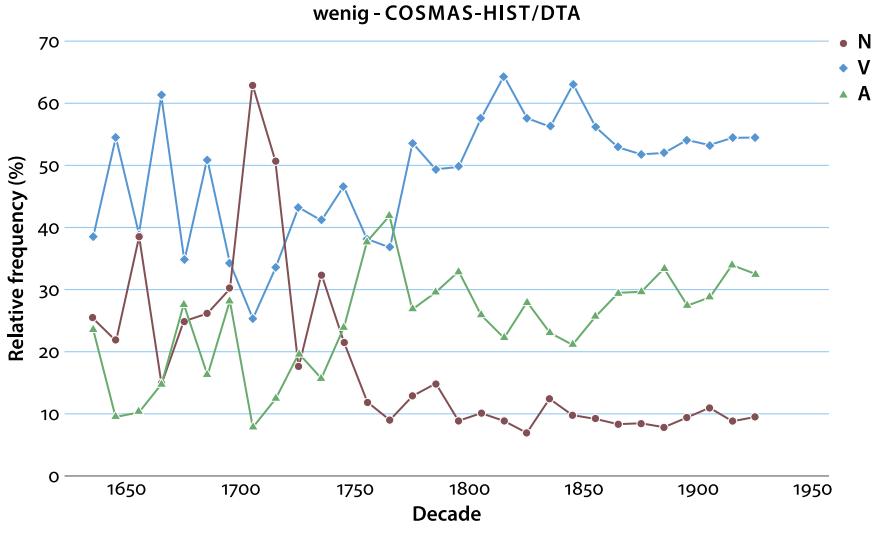


Figure 3. Relative frequencies of nouns, verbs, and adjectives modified by *ein wenig* in proportion to the sum total of items modified by *ein wenig* across the decades of the joint DTA/DeReKo-HIST database

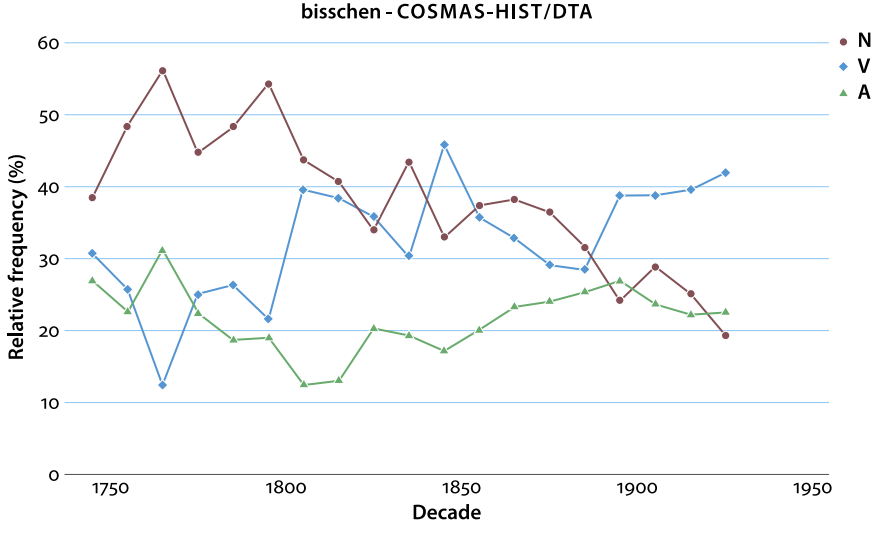


Figure 4. Relative frequencies of nouns, verbs, and adjectives modified by (*ein*) *bisschen* in proportion to the sum total of items modified by (*ein*) *bisschen* across the decades of the joint DTA/DeReKo-HIST database

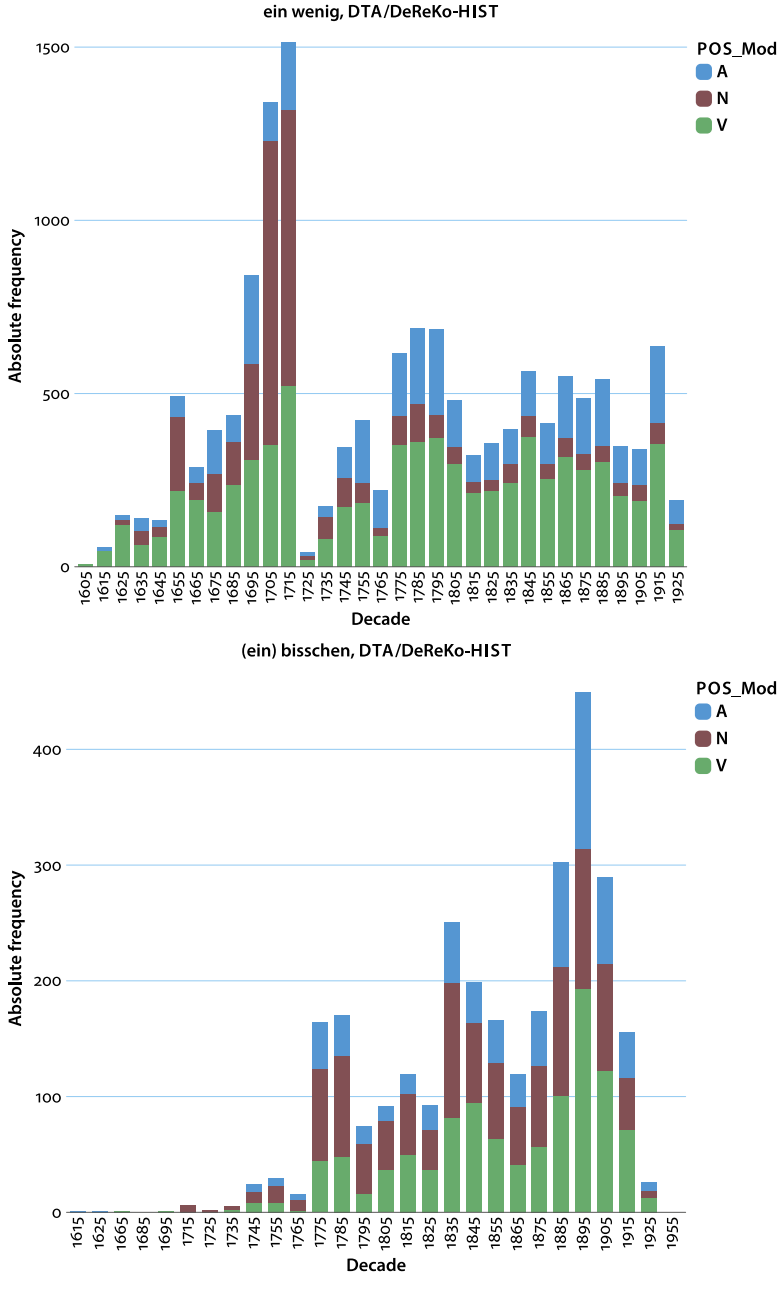


Figure 5. Absolute frequencies of nouns, verbs, and adjectives modified by *ein wenig* and (*ein*) *bisschen* in the joint DTA/DeReKo-HIST database

from a literal use to partitive, quantifier, and degree-modifier uses. If the concrete use of *ein bisschen*, i.e. 'a bite of', represents the starting point for its subsequent grammaticalization into a quantifier and degree modifier, it can be expected that the construction is, at first, mostly constrained to modifying nouns. In addition, we can assume that the first nouns to be modified by *ein bisschen* were concrete nouns.

To test this hypothesis, all nouns in the DTA/DeReKo-HIST joint database were annotated for their concreteness, adopting a coarse-grained distinction between 'concrete' and 'abstract' items. Figure 6 shows that abstract nouns become more frequent in both constructions. In both cases, the increase is statistically significant according to a Kendall's tau test (*bisschen*: $\tau = 0.37$, $z = 2.42$, $p < 0.05$; *wenig*: $\tau = 0.64$, $T = 356$, $p < 0.001$). Surprisingly, the increase is much clearer in the case of the older construction *ein wenig*, while *ein bisschen* features abstract *modificata* from the beginning. The very low relative frequency of abstract nouns in the [*ein wenig* N] construction before 1700 might be an artefact of the corpus structure, as this period contains cooking recipes in which *ein wenig* + concrete noun occurs very frequently (e.g. *ein wenig Wein* 'a little wine' etc.). However, the increase is quite robust over the entire period in question. This observation lends support to the hypothesis that the existence of a paradigmatic relation between *ein wenig* and *ein bisschen* can be assumed from the earliest stages of the younger construction. When the [*ein bisschen* X] construction grammaticalized, the modification of abstract nouns (and other parts of speech, for that matter) with *ein wenig* was already widespread. This pattern could be easily adopted for the formally similar [*ein bisschen* X] construction. This in turn is in line with the idea that a joint meso-construction can be assumed for the German quantifier/degree-modifier constructions, which also includes the older variant *ein lützel*. The subordinate constructions belonging to this meso-construction are connected via paradigmatic relations, and their diachronic development is very likely to be a coevolutionary one. This idea will be developed further in Section 8.

In sum, then, the data lend support to a grammaticalization scenario similar to the one discussed by Traugott (e.g. 2008a) for English. In the case of *bisschen*, the modified nouns tend to be concrete nouns at first. Gradually, however, the distribution of nouns, verbs, and adjectives roughly aligns to that of *ein wenig*. In addition, the patterns of 'co-evolution' between *ein bisschen* on the one hand and *ein wenig* on the other lend support to the idea that both constructions share a common more abstract schema, just like their English counterparts can be seen as instantiations of an [NP *of* NP] meso-construction.

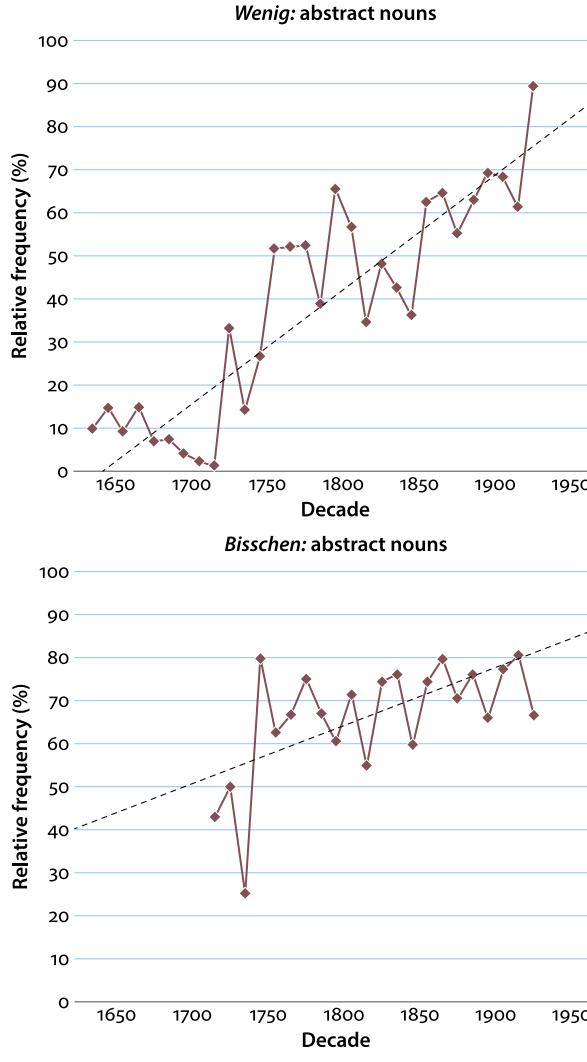


Figure 6. Relative frequencies of abstract nouns modified by *ein wenig* and (*ein*) *bisschen* in the joint DTA/DeReKo-HIST database (in relation to the total number of nouns modified by the respective construction)

6.3 Determiner variation

As pointed out above, the use of determiners varies in the case of (*ein*) *bisschen*, while *wenig* usually cannot be used as a quantifier or degree modifier without the indefinite determiner *ein*. However, given the hypothesis that *bisschen* gradually assimilates to the general [*ein* N X] pattern of the quantifier/degree-modifier meso-construction, we can expect the variation between different determiners to

decrease over time. In fact, this is exactly what we find in the data (see Figure 7). The use of *ein* together with *bisschen*, as opposed to definite articles or possessive and demonstrative determiners (e.g. *das*, *mein*, *dieses* respectively), increases significantly over the time period covered by DTA and DeReKo-HIST (Kendall's $\tau = 0.55$, $T = 163$, $p < 0.001$). In more general terms, this finding translates into a fixing or freezing of form that is frequently seen as characteristic of grammaticalization, especially so in the "grammaticalization-as-reduction" approach.

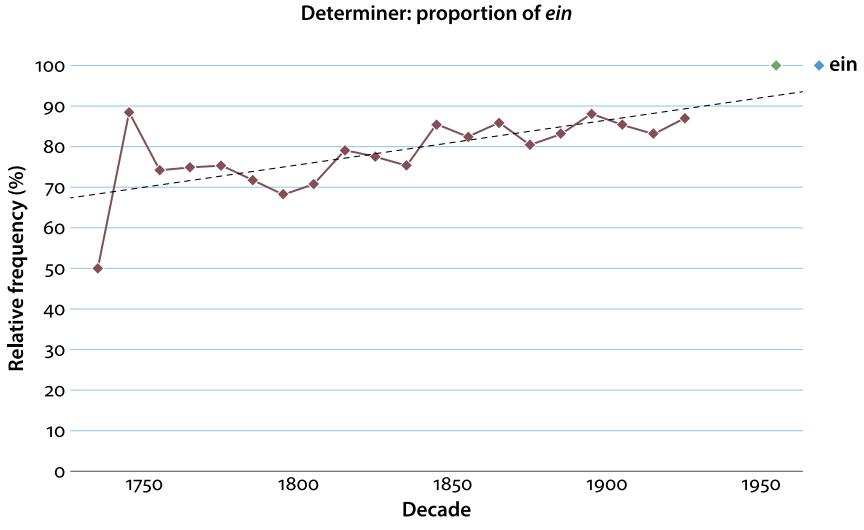


Figure 7. Relative frequency of *bisschen* used with the indefinite article *ein* in proportion to the total number of attestations of *bisschen* in the joint DTA/DeReKo-HIST database

7. Bridging contexts and motivations for innovation

When adopting a more qualitative perspective, the corpus data can shed light on how specific usage events provide potential for diachronic change. Several constructs in the sample are ambiguous, thus illustrating bridging contexts (Heine, 2002) for formal and functional reanalysis.[4] The appearance of uses such as (15)

4. Note that we understand *reanalysis* as a process of restructuring that is neither necessarily abrupt nor confined to children's language acquisition. Associations such as 'abruptness of change' and 'discrete category boundaries' have traditionally been linked to reanalysis; however, a notion of reanalysis burdened with these additional theoretical associations is hard to reconcile with usage-based grammaticalization theory (cf. Haspelmath, 1998; De Smet, 2009, 2014).

suggests that a syntactic reanalysis of the kind described by Traugott (cf. (4) above) did occur in German as well.

(15) *Hätt' ich's doch nicht gedacht, daß in **der bißchen Neige** meines Lebens noch etwas Bittrers wäre als Tod!*
'I wouldn't have thought that, in the FEM bit NEUT of remainder FEM of my life, something was more bitter than death!' (1776, DereKo-HIST)

In this sentence, the definite article *der* shows gender agreement, not with *bißchen*, but with the following noun *Neige* 'remainder'. This clearly indicates a reconfiguration of constituent structure such that *bisschen* was demoted from head noun to modifier. Given that uses without any preceding determiner at all have also become acceptable in colloquial German (cf. *Ich lese bisschen*, Example (14) above), there is little doubt that *bisschen* ceased to be an ordinary count noun heading a noun phrase. Changes of this sort are predicted by Hopper's (1991) grammaticalization principle of decategorialization.

Other constructs represent bridges between established and innovative functions. For illustration, consider the attested sentences in (16) to (21), all of which allow for multiple readings.

(16) *Mariechen – sagte ich – so iß doch nur **en bischen!***
'Mariechen – I said – (please) eat just a bit(e)!' (1801, DeReKo-HIST)

(17) *Ich habe aber doch Mittel gefunden, dem Kinde einigemal **ein Bißchen** Brod zuzustecken.*
'But I found means to slip the child a bit of bread a few times.'
(1784, DeReKo-Hist)

(18) *Und von dieser [Versuchung] muß ich dir auch noch **ein Bischen** erzählen, mein Sohn!*
'And of this [temptation] I have to tell you a bit, my son!'
(1788, DeReKo-HIST)

(19) *… wo man nur während des Karnevals **ein bißchen** Cancan tanzt*
'… where they dance a bit of Cancan / dance the Cancan a bit only during carnival' (1840, DeReKo-HIST)

(20) *Ist etwa die Brühe zu dünne, magst du **ein wenig** in Butter braun geröstet Mehl hinein rühren;*
'If the broth is too watery, you may stir in a little flour that was browned in butter / stir in flour that was browned in butter a bit' (1715, DTA)

(21) *Ich wollte es noch **ein wenig** geheim halten, aber sie haben sich selber verrathen*
'I wanted to keep it secret a little (while), but they gave themselves away
(1844, DeReKo-HIST)

The first two examples, occurring in the context of ingestion, are very likely to instantiate literal uses of the noun *Bisschen*, but there is still some slight uncertainty about its actual function, which could range from the pre-partitive to the quantifier type. Similarly, Example (18) is ambiguous between a partitive and a degree-modifier reading. While the speaker may have had in mind a conservative partitive use, for the hearer or reader a degree-modifier interpretation is equally likely, especially given the adjacency of *Bischen* to the verb *erzählen* 'tell' and its distance to the preceding PP *von dieser* 'of this'. Examples (19) and (20) allow for both quantifier and degree-modifier readings: the elements modified can be either the nouns (*Cancan*, *Mehl* 'flour') or the verbs (*tanzen* 'dance', *geröstet* 'roasted'). Sentence (21) clearly contains a degree-modifier use of *ein wenig*, but there is some ambiguity as to whether the modifier *ein wenig* is more closely linked to the adjectival part or the verb part of *geheim halten* 'keep secret'. Such minor indeterminacies in constituency relate to some of the finer semantic nuances of *ein wenig* and *ein bisschen*. When combined with adjectives, these German modifiers specify intensities and degrees of properties. When modifying verbs, however, they often do not express intensities of activities but rather senses that have to do with (short) durations of time. Systematic polysemies between older and newer meanings have remained to the present day. Consequently, the diachronic grammaticalization path of these constructions is largely mirrored in their synchronic polyfunctionality.

Both hearer-centred and speaker-centred motivations are relevant to functional expansion, structural reinterpretation and other kinds of linguistic innovation. Hearers are generally inclined to draw rich contextual inferences when interpreting utterances. Pragmatic inferencing, usually grounded in conceptual metonymy (e.g. Panther & Thornburg, 2003), is a key factor in various approaches to meaning change, such as Traugott & Dasher's (2002) Invited Inferencing Theory of Semantic Change, Heine et al.'s (1991) notion of context-induced reinterpretation, and Croft's (2000) metanalysis. Speakers may exploit invited inferences and produce innovative constructs, motivated by the maxim of expressiveness or 'extravagance' (Haspelmath, 1999). Through this combination of production- and comprehension-oriented factors, non-conventional or ambiguous constructs like the ones exemplified above can, if experienced repeatedly, alter the cognitive representation of the construction they instantiate.

8. Layering, attractor sets and constructional levels

As *ein lützel*, *ein wenig* and *ein bisschen* coexisted in Early New High German (ca. 1350–1650), the three micro-constructions entered into a paradigmatic rela-

tion. What the quantitative data on *ein wenig* and *ein bisschen* suggest is that the younger micro-constructions gradually converged with the older ones, both functionally and formally. This happened despite the fact that the individual constructions originated from sources that were formally rather different, given that *bisschen* started as a noun and *wenig* as an adjective, for instance (cf. Section 4). The convergence of the usage patterns of the *ein bisschen* construction with those of *ein wenig* indicate that the grammaticalization of *ein bisschen* was not independent but strongly assisted by paradigmatic associations to the more established *ein wenig* construction. Likewise, it can be assumed that the development of *ein wenig* was influenced by the yet older *ein lützel* construction.

Links in constructional networks point to possible directions of functional expansion. Language users, when processing non-conventional and ambiguous constructs that do not easily match an existing micro-construction, can have recourse to other constructions that are deemed similar. In their diachronic construction grammar framework, Traugott & Trousdale (2013, p. 57) assume that

> [w]hen the hearer attempts to match a construct with an existing part of the constructional network and fails to do so because there is no existing micro-construction that fully sanctions the construct, there is mismatch. The best the hearer can do is create a link to align the meaning or the form of the construct with the meaning or form of other extant (sub)schemas in the network. This is done based on the discourse/pragmatic properties associated with both the (new) construct and the (existing) constructional subschema [i.e. meso-construction].

More specifically, encountering a novel *ein bisschen* construct in a bridging context that also makes an innovative degree-modifier interpretation plausible can arouse new associations to the *ein wenig* construction, which readily sanctions degree-modifier uses. When parsing subsequent usage events with [*ein bisschen* X], language users may co-activate [*ein wenig* X], and the link between both micro-constructions becomes strengthened. In addition, nodes higher up in the constructional network might be formed and strengthened, as will be discussed further below. The marked or extravagant status of early degree-modifier uses of *ein bisschen* formed analogously to *ein wenig* diminished over time through the habituating effect of repeated experience (cf. Haiman, 1994). In order to accommodate emerging degree-modifier constructs, the *ein bisschen* construction was incited to generalize in meaning. As a result, it successfully expanded its functional range.

As in English, there is a whole family of quantifier/degree-modifier constructions in present-day German. They form a layered domain of grammar (cf. Hopper, 1991). Further innovative modifiers have gained ground more recently,

such as *ein Quäntchen* 'a quantum$_{DIM}$', *eine Prise* 'a pinch', *ein Funken* 'a spark' (cf. *ein Fünklein*, discussed in Section 4), *ein Tick* 'a tick' and *eine Idee* 'an idea'.

(22) *Immerhin, auch in dieser schönen neuen Arbeitswelt ist noch Platz für* ***ein Quäntchen*** *persönlicher Idylle.*
'At least, even in this nice new world of work, there is room for a bit (lit.: quantum$_{DIM}$) of a personal idyll' (1999, *Die Zeit*)

(23) ***Eine Prise*** *Liberalismus wird dem Land guttun*
'a pinch of liberalism will do the country good' (2012, *Die Zeit*)

(24) *Aber wenn die Bundesregierung ihren eigenen Bericht ernst nimmt und noch über* ***einen Funken*** *Anstand verfügt, muß sie Stahl jetzt wieder in sein Amt einsetzen.*
'But if the Federal Government takes their own report seriously and still has a spark of decency, they must put Stahl back into his office.' (1993, D*ie Zeit*)

(25) *Joachim Löw sah, wie seine Spieler meist* ***einen Tick*** *schneller waren als ihre Gegner*
'Joachim Löw saw that his players were usually a bit faster than their opponents' (2008, *Die Zeit*)

(26) *Das klang dem anderen Gast des Abends dann doch* ***eine Idee*** *zu pessimistisch.*
'To the other guest of the evening, this sounded a bit (lit.: idea) too pessimistic.' (2008, *Die Zeit*)

Compared to *ein wenig* and *ein bisschen*, younger modifiers like the ones exemplified above (and the ones mentioned in Section 4) are used considerably less frequently and still have far more syntagmatic restrictions. For example, *ein Tick* and *eine Idee* combine almost exclusively with adjectives, and, more specifically, with graded adjectives, as in (25) and (26). The quantifiers *ein Quäntchen* and *ein Funken*, on the other hand, are still strongly associated with a few individual collocates, forming common expressions such as *ein Quäntchen Glück* 'a quantum$_{DIM}$ of luck' or *ein Funken Hoffnung* 'a spark of hope'. This is reminiscent of Early New High German uses of *ein bisschen* when the expression exhibited a clear preference for 'edible' collocates that is no longer prominent in today's German. Each modifier has a distinct history that strongly depends on its lexical source, and there are effects of persistence (Hopper, 1991). Clearly, these individual histories would merit further investigation. For instance, *Idee* (which is first attested in DTA and DeReKo in 1855 in a degree-modifier construction: *eine Idee stärger* 'an idea stronger') seems like a fairly atypical source for a degree modifier at first glance. On second thought, however, several explanations readily suggest themselves. Firstly, *Idee* refers to something that is construed as an individuated entity –

this is supported by the fact that it occurs in the pluralized form (which suggests a count noun construal) in around 40% of all cases attested in the DTA. Secondly, while ideas can obviously be big or small, some sense of smallness might be attached to the word via iconicity and sound symbolism, as *Idee* is a one-syllable word (iconicity: shortness implicates smallness) and front vowels tend to be associated with smallness (see e.g. Preziosi & Coane, 2017; Hartmann et al., forthcoming).

On the other hand, also clear tendencies of diachronic convergence can be found in this family of quantifier/degree-modifier constructions, as the quantitative corpus analyses on *ein bisschen* and *ein wenig* have revealed. For this convergence and addition of constructions to happen, a number of forces had to operate in conjunction with each other. First, we take the described development to lend support to the hypothesis that a higher-level quantifier/degree-modifier schema can be assumed for the micro-constructions under discussion. From ca. 1500 onwards, with *ein lützel*, *ein wenig* and *ein bisschen* coexisting and reinforcing each other, a common abstract schema became more and more firmly entrenched. Figure 8 is an attempt to visualize this schema or meso-construction.

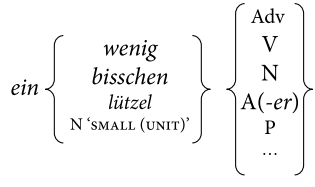


Figure 8. Mid-level quantifier/degree-modifier schema for (Early) New High German. (The suffix *-er* on adjectives marks their comparative form; some younger degree-modifier constructions can only be used with this form, e.g. *einen Tick schneller* 'a tick faster', but **einen Tick schnell* 'a tick fast')

Crucially, the postulated mid-level schema has a specific form (including *ein* 'a') that makes it distinct from other schemas covering structurally different quantifiers and degree modifiers such as the one-word items *etwas* 'somewhat', *leicht* 'slightly', *sehr* 'very' and *viel* 'much'. That is, whereas *ein bisschen* and *ein wenig* are sanctioned by the same meso-construction (Figure 8), modifiers that do not share a structural pattern, like *ein bisschen* and *etwas*, are assumed to be linked only at an even higher level of abstraction, namely the degree-modifier macro-construction. Second, the addition of *ein Funken*, *eine Idee*, *ein Quäntchen*, *ein Tick* and other innovative modifiers to the schema shown in Figure 8 would have been very unlikely if there were not a few well-established micro-constructions that can serve as attractors for new schema members via analogical pattern matching. Analogy, as pointed out by Fischer (2010, 2011, 2013), must be acknowledged to

operate pervasively as a fundamental mechanism of language change from a processing point of view. In contemporary German, *ein wenig* and *ein bisschen* are interchangeable in most contexts, and they are highly productive and in frequent use. A language user who regularly experiences [*ein bisschen* X] and [*ein wenig* X] as functionally equivalent quantifier/degree-modifier constructions may develop some degree of (tacit) awareness of the semantic commonality of the lexical elements *Bisschen* and *wenig*, which both have 'small (unit)' as central meaning component. The semantic component 'small unit' is shared by the lexical sources for [*ein Funken* X], [*eine Idee* X], [*ein Tick* X], [*ein Quäntchen* X] and [*eine Prise* X], all of which express downtoning meanings accordingly. It thus appears that *ein bisschen* and *ein wenig*, thanks to their respectable discourse frequencies, productivity and interchangeability, acted as models for the emergence of novel modifiers and/or for their functional expansion.

This assumption is in keeping with data on Spanish binominal quantifiers discussed by Verveckken (2012). She argues that the old highly frequent and highly grammaticalized construction *un montón de* 'a heap of' attracted new members to the quantifying [N1 *de* N2] schema via analogical processes that are semantically or conceptually driven. *Un montón de* is said to have served as a prototype in the emergence of a number of younger low-frequency quantifiers such as *un hatajo de* 'a herd of', *una letanía de* 'a litany of' and *un racimo de* 'a bunch of', which still exhibit strong formal and conceptual persistence, however. Besides paradigmatization, layering and persistence, Verveckken identifies several other concomitant grammaticalization features in her case study on Spanish, such as rebracketing (or syntactic reanalysis), decategorialization and context expansion. As pointed out above, all of these features apply to the development of the German constructions as well.

There is an empirical fact that complicates a unitary analysis of the German quantifier/degree-modifier constructions at hand, though. Not all of the younger modifiers have both quantifier and degree-modifier uses. *Ein Quäntchen* does serve both functions (e.g. *ein Quäntchen Effizienz* 'a quantum of efficiency' or *ein Quäntchen effizienter* 'a bit more efficient'), just like *ein bisschen* and *ein wenig*. However, *ein Tick* and *eine Idee* virtually always degree-modify graded adjectives; quantifier uses of *ein Tick* or *eine Idee* are very exceptional in (standard) German. *Eine Prise* and *ein Funken*, on the other hand, can in general only be used to quantify nouns. Micro-constructions with quantifier uses but no degree-modifier uses might simply be said to have not (yet) gone this step in the grammaticalization path suggested by Traugott for English (pre-partitive > partitive > quantifier > degree modifier > free adverb). More problematic for the idea of a universal cline are *ein Tick* and *eine Idee*. They would have to be said to have directly jumped to degree-modifier functions, skipping the "less grammaticalized" quantifier stage.

Interestingly, similar grammaticalization scenarios with unattested (supposedly) intermediate stages have also been observed by De Clerck & Brems (2016) in their related study on English size nouns like *mass*(*es*), *bunch* and *heap*(*s*). In Section 7, it has been illustrated that some contexts of use in German (e.g. Example (18)) bear ambiguity bridging from partitive directly to degree-modifier functions, which suggests that the chronology of Traugott's cline need not be interpreted too rigidly, at least for German. Differences in the grammaticalization paths of the individual micro-constructions, including the 'leap' of *ein Tick* and *eine Idee* to the degree-modifier stage, demand for explanations. In line with a possibility that De Clerck & Brems (2016, p. 175) consider in their case study, our interpretation is as follows: the younger German constructions did not necessarily have to follow the original chronology of functional expansion because they could instead join an extant schema via analogy, with older micro-constructions serving as attractor sets for this schema.

A closer examination of the usage patterns in this family of constructions highlights the individual trajectories and properties of each micro-construction. At the same time, these formally and functionally highly similar micro-constructions have also been influencing and reinforcing each other in their entrenchment and use. Given the fact that some of the younger, less established expressions discussed here have only quantifier uses and others only degree-modifier uses, it appears that these two functions have in fact continued to be represented by two separate (sub)schemas. What is more, while beyond the scope of the present paper, in-depth analyses of the younger modifiers will certainly underline the need to postulate further, even more specific subschemas in order to capture individual usage constraints. For example, within the degree-modifier meso-construction, there seems to be a firmly entrenched subschema specifically with graded adjectives; this subschema appears to sanction the majority of uses of the modifiers *ein Tick* and *eine Idee* (e.g. Examples (25) and (26)). Pronounced individual differences between micro-constructions may pose a challenge to our account that assigns the older micro-constructions a decisive role as models for the emergence or change of younger constructions. It is, however, important to reemphasize that the by now highly frequent and unconstrained *ein bisschen* and *ein wenig* participate in all subschemas that the younger low-frequent modifiers mentioned above are currently limited to. This includes, for instance, the graded-adjective subschema [*ein* X A-*er*] (as indicated in Figure 8), which is instantiated by more than 20% of all adjective combinations of *ein bisschen* and *ein wenig* in our historical corpus data. All in all, there is thus not only a heavy overlap in structure but also some overlap in the members shared by the quantifier meso-construction and the degree-modifier meso-construction as well as by their subschemas. These overlaps both in schema structure and in (sub)schema

members facilitate constructional change and, in particular, the context expansion of individual micro-constructions. The suggested constellation of network links between constructional levels promotes realignments of micro-constructions, allowing, for instance, the younger *ein Quäntchen* to expand and participate in the degree-modifier schema in addition to the related quantifier schema.

9. Conclusion and outlook

In response to the research questions raised in the Introduction, three generalizations can be drawn from this case study. First, grammaticalization paths are once more shown to be valid cross-linguistically. The specific trajectory that Traugott observed in the evolution of multi-word degree modifiers in English seems to apply to the German modifiers as well. This holds true especially for the older German constructions *ein bisschen* and *ein wenig*; some of the younger low-frequency modifiers in the family, in particular *ein Tick* and *eine Idee*, did not pass through the stages of the grammaticalization path in successive order, apparently because they could directly join a strengthened degree-modifier subschema by analogy. Second, grammaticalization is neither solely expansive nor solely reductive in nature, but, as our large sample of historical usage data illustrates, both kinds of change operate in the co-evolution of function and form. The development of the German degree modifiers involves reduction in individual formal variation and expansion in their range of use. From a construction grammar point of view, the linguistic units affected by these reductive and expansive changes are form–meaning pairings whose substantive core elements became more tightly bonded and whose slots became increasingly schematic and productive. Reduction seems to be confined basically to the level of micro-constructions. Expansion, on this constructional level, manifested itself in the construction-final slot (e.g. [*ein bisschen* X]); on the meso-constructional level, another slot following *ein* ([*ein* N 'small unit' X]) has been emerging, showing some signs of expandability and incipient productivity. The formation and entrenchment of such a mid-level schema is tied to the coexistence and supportive interaction of multiple sufficiently similar micro-constructions. More specifically, we have argued that older, well-established micro-constructions served as attractor sets promoting the analogical addition of new members to the mid-level schema they instantiate. These observations, thirdly, stress the crucial role that network links between constructions can play in grammaticalization processes, thus providing valuable insights into mechanisms and motivations driving grammaticalization and language change in general.

Further questions prompted by the present investigation remain open for future research. We have depicted the interaction of multiple constructions mainly as a case of mutual support facilitating the grammaticalization of individual micro-constructions as well as the strengthening of mid-level schemas. In other approaches to the grammaticalization of interrelated constructions, most notably in variationist studies (see, e.g., Poplack, 2011), interactions between constructions tend to be depicted as competition rather than mutual support. It remains to be clarified how the construction grammar view of supporting networks can be reconciled with the view of competing variants, which also plays a key role in current constructionist approaches (cf. Barðdal, 2008; Delorge et al., 2014; among others).

Other open questions concern the refinement of usage-based explanations of grammaticalization drawing on frequency effects (e.g. Bybee, 2003a, 2003b, 2010). Can low-frequent expressions grammaticalize only when assisted by strong paradigmatic associations to other, more frequent grammatical micro- and meso-constructions? The studies by Hoffmann (2004), Brems (2007, 2011) and Verveckken (2012) point in this direction, and so does the study at hand. Is it plausible to postulate frequency thresholds that must be satisfied for grammaticalization to occur in the absence of such a supporting network of paradigmatic associations? If so, these frequency thresholds must certainly be specific to different grammatical domains (e.g. auxiliaries, prepositions, degree modifiers), and they probably cannot be based solely on token frequency but should take into account other notions such as conceptual frequency (Hoffmann, 2004). Furthermore, more research should be devoted to assessing which type–token ratios and other factors are likely to lead to the strengthening of higher-level schemas in addition to the entrenchment and routinization of individual micro-constructions (cf. the so-called upward strengthening hypothesis by Hilpert, 2015). Converging evidence from both corpus-linguistic studies and psycholinguistic experimentation is required in order to find sound answers to these questions.

Acknowledgements

We would like to thank the editors of the present volume as well as the anonymous reviewers for their helpful comments and suggestions on a previous draft of this paper. We are also grateful to the audiences at the workshop "Grammaticalization meets Construction Grammar" in Gothenburg (October 2015) and at the SLE conference in Naples (August–September 2016) for stimulating feedback.

Corpora and sources

[DeReKo] German Reference Corpus (Deutsches Referenzkorpus), available via COSMAS II, https://cosmas2.ids-mannheim.de/cosmas2-web/
[DTA] German Text Archive (Deutsches Textarchiv), http://www.deutschestextarchiv.de/
[DWDS] Digital Dictionary of the German Language (Digitales Wörterbuch der deutschen Sprache), http://dwds.de/
[Internet Archive] www.archive.org
[MED] Middle English Dictionary, University of Michigan, http://quod.lib.umich.edu/m/med/
[MHDBDB] Middle High German Conceptual Database (Mittelhochdeutsche Begriffsdatenbank), http://mhdbdb.sbg.ac.at/.
[PCEEC] Parsed Corpus of Early English Correspondence, compiled by T. Nevalainen, H. Raumolin-Brunberg, J. Keränen, M. Nevala, A. Nurmi, & M. Palander-Collin (2006), distributed through the Oxford Text Archive

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511801686
Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage-based models of language*. Stanford, CA: CLSI Publications.
Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8
Brems, L. (2007). The grammaticalization of small size nouns: Reconsidering frequency and analogy. *Journal of English Linguistics*, 35(4), 293–324. https://doi.org/10.1177/0075424207307597
Brems, L. (2011). *Layering of size and type noun constructions in English*. Berlin: De Gruyter. https://doi.org/10.1515/9783110252927
Brems, L. (2012). The establishment of quantifier constructions for size nouns: A diachronic case study of *heap(s)* and *lot(s)* . *Journal of Historical Pragmatics*, 13(2), 202–231. https://doi.org/10.1075/jhp.13.2.02bre
Bybee, J. L. (2003a). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure*. Volume 2. Mahwah, NJ: Erlbaum.
Bybee, J. L. (2003b). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph, & R. Janda (Eds.), *The handbook of historical linguistics* (pp. 602–623). Malden: Blackwell. https://doi.org/10.1002/9780470756393.ch19
Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526
Bybee, J. L., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow: Pearson Longman.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

De Clerck, B., & Brems, L. (2016). Size nouns matter: a closer look at *mass(es) of* and extended uses of SNs. *Language Sciences*, 53, 160–176. https://doi.org/10.1016/j.langsci.2015.05.007

De Clerck, B., & Colleman, T. (2013). From noun to intensifier: *massa* and *massa's* in Flemish varieties of Dutch. *Language Sciences*, 36, 147–160. https://doi.org/10.1016/j.langsci.2012.04.005

Delorge, M., Plevoets, K., & Colleman, T. (2014). Competing "transfer" constructions in Dutch. The case of ont-verbs. In D. Glynn, & J. A. Robinson (Eds.), *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy* (pp. 39–60). Amsterdam: John Benjamins. https://doi.org/10.1075/hcp.43.02del

De Smet, H. (2009). Analyzing reanalysis. *Lingua*, 119, 1728–1755. https://doi.org/10.1016/j.lingua.2009.03.001

De Smet, H. (2014). Does innovation need reanalysis? In E. Coussé, & F. von Mengden (Eds.), *Usage-based approaches to language change* (pp. 23–48). Amsterdam: John Benjamins.

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25, 108–127. https://doi.org/10.1016/j.newideapsych.2007.02.002

Diessel, H. (2015). Usage-based construction grammar. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 296–322). Berlin: De Gruyter. https://doi.org/10.1515/9783110292022-015

Diessel, H., & Hilpert, M. (2016). Frequency effects in grammar. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. New York: Oxford University Press.

Durrell, M., Ensslin, A., & Bennett, P. (2007). The GerManC Project. *Sprache und Datenverarbeitung*, 31, 71–80.

Fischer, O. (2010). An iconic analogical approach to grammaticalization. In J. Conradie, R. Johl, M. Beukes, O. Fischer, & C. Ljungberg (Eds.), *Signergy* (pp. 279–298). Amsterdam: John Benjamins. https://doi.org/10.1075/ill.9.17fis

Fischer, O. (2011). Grammaticalization as analogically driven change? In H. Narrog, & B. Heine (Eds.), *Oxford handbook of grammaticalization* (pp. 31–42). Oxford: Oxford University Press.

Fischer, O. (2013). An inquiry into unidirectionality as a foundational element of grammaticalization: On the role played by analogy and the synchronic grammar system in processes of language change. *Studies in Language*, 37(3), 515–533. https://doi.org/10.1075/sl.37.3.03fis

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Haiman, J. (1994). Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on grammaticalization* (pp. 3–28). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.109.07hai

Hartmann, S., Pleyer, M., Winters, J., & Zlatev, J. (Eds.). (forthcoming). Interaction and iconicity in the evolution of language. *Special issue of Interaction Studies* .

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language*, 22(2), 315–351. https://doi.org/10.1075/sl.22.2.03has

Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics*, 37(6), 1043–1068. https://doi.org/10.1515/ling.37.6.1043

Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 83–101). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Hilpert, M. (2014). *Construction grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hilpert, M. (2015). From *hand-carved* to *computer-based*: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics*, 26(1), 113–147. https://doi.org/10.1515/cog-2014-0001

Hilpert, M., & Gries, S. T. (2009). Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language Acquisition. *Literary and Linguistic Computing*, 24(4), 385–401. https://doi.org/10.1093/llc/fqn012

Himmelmann, N. P. (2004). Lexicalization and grammaticization: Opposite or orthogonal? In W. Bisang, N. P. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization?* (pp. 21–42). Berlin: De Gruyter.

Hoffmann, S. (2004). Are low-frequency complex prepositions grammaticalized? On the limits of corpus data – and the importance of intuition. In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 171–210). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.13.09hof

Hopper, P. J. (1987). Emergent grammar. *Berkeley Linguistics Society*, 13, 139–157.

Hopper, P. J. (1991). On some principles of grammaticalization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to grammaticalization*, Volume 1 (pp. 17–36). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.19.1.04hop

Hopper, P. J., & Traugott E. C. (2003). *Grammaticalization*. Second edition. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Kluge, F. (2012). *Etymologisches Wörterbuch der deutschen Sprache: Bearbeitet von Elmar Seebold*. 25th edition. Berlin: De Gruyter. https://doi.org/10.1515/9783110223651

Krug, M. G. (2000). *Emerging English modals: A corpus-based study of grammaticalization*. Berlin: De Gruyter. https://doi.org/10.1515/9783110820980

Krug, M. G. (2003). Frequency as a determinant in grammatical variation and change. In G. Rohdenburg, & B. Mondorf (Eds.), *Determinants of grammatical variation in English* (pp. 7–67). Berlin: De Gruyter. https://doi.org/10.1515/9783110900019.7

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). Stanford, CA: CLSI Publications.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331967.001.0001

Lehmann, C. (2015). *Thoughts on grammaticalization*. 3rd edition. Berlin: Language Science Press.

Lehmann, C. (2016). *Grammaticalization and automation*. Paper presented at the 23rd LIPP-Symposium: Grammatikalisierung in interdisziplinärer Perspektive, Munich, Germany. http://www.christianlehmann.eu/publ/lehmann_automation.pdf

Panther, K.-U., & Thornburg, L. (Eds.). (2003). *Metonymy and pragmatic inferencing*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.113

Pfeifer, W. (1993). *Etymologisches Wörterbuch des Deutschen*. Berlin: Akademie-Verlag.

Poplack, S. (2011). Grammaticalization and linguistic variation. In H. Narrog, & B. Heine (Eds.), *Oxford Handbook of Grammaticalization* (pp. 209–224). Oxford: Oxford University Press.

Preziosi, M. A., & Coane, J. H. (2017). Remembering that big things sound big: Sound symbolism and associative memory. *Cognitive Research: Principles and Implications*, 2(10). https://doi.org/10.1186/s41235-016-0047-y.

Sweetser, E. (1988). Grammaticalization and semantic bleaching. In S. Axmaker, A. Jaisser, & H. Singmaster (Eds.), *Berkeley Linguistics Society 14: General session and parasession on grammaticalization* (pp. 389–405). Berkeley: Berkeley Linguistics Society.

Tabor, W., & Traugott, E. C. (1998). Structural scope expansion and grammaticalization. In A. G. Ramat, & P. J. Hopper (Eds.), *The limits of grammaticalization* (pp. 227–270). Amsterdam: John Benjamins.

Tiefenbach, H. (1987). -chen und -lein. Überlegungen zu Problemen des sprachgeographischen Befundes und seiner sprachhistorischen Deutung. *Zeitschrift für Dialektologie und Linguistik*, 54(1), 2–27.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Traugott, E. C. (1988). Pragmatic strengthening and grammaticalization. In S. Axmaker, A. Jaisser, & H. Singmaster (Eds.), *Berkeley Linguistics Society 14: General session and parasession on grammaticalization* (pp. 406–416). Berkeley: Berkeley Linguistics Society.

Traugott, E. C. (2007). The concepts of constructional mismatch and type-shifting from the perspective of grammaticalization. *Cognitive Linguistics*, 18(4), 523–557. https://doi.org/10.1515/COG.2007.027

Traugott, E. C. (2008a). Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In R. Eckhart, G. Jäger, & T. Veenstra (Eds.), *Variation, selection, development* (pp. 219–250). Berlin: De Gruyter.

Traugott, E. C. (2008b). The Grammaticalization of NP of NP Patterns. In A. Bergs, & G. Diewald (Eds.), *Constructions and language change* (pp. 23–45). Berlin: De Gruyter.

Traugott, E. C. (2010). Grammaticalization. In S. Luraghi, & V. Bubenik (Eds.), *Continuum companion to historical linguistics* (pp. 269–283). London: Continuum Press.

Traugott, E. C. (2014). Towards a constructional framework for studying language change. *Cognitive Linguistic Studies*, 1(1), 3–21. https://doi.org/10.1075/cogls.1.1.01tra

Traugott, E. C., & Dasher, R. B. (2002). *Regularities in semantic change*. Cambridge: Cambridge University Press.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Trousdale, G., & Norde, M. (2013). Degrammaticalization and constructionalization: Two case studies. *Language Sciences*, 36, 32–46. https://doi.org/10.1016/j.langsci.2012.03.018

Verveckken, K. D. (2012). Towards a constructional account of high and low frequency binominal quantifiers in Spanish. *Cognitive Linguistics*, 23(2), 421–478. https://doi.org/10.1515/cog-2012-0013

Verveckken, K. D. (2015). *Binominal quantifiers in Spanish: conceptually-driven analogy in diachrony and synchrony*. Berlin: De Gruyter. https://doi.org/10.1515/9783110406733

Type frequency, productivity and schematicity in the evolution of the Latin *secundum* NP construction

Caterina Guardamagna
University of Liverpool

This paper adopts a constructional approach as a heuristic to explore the changes undergone by Latin *secundum* NP 'according to NP', considering both its frequency and its network organisation. My data demonstrate a significant syntactic and host-class expansion of *secundum* NP between the years BC 106 – AD 17 and a support to Barðdal's (2008) model of productivity. The vocabulary growth curve for the construction also suggests that the productivity of *secundum* NP remains high in all the periods considered. Finally, the discussion highlights the role played by the social context in stimulating some of the changes observed, which is crucial in order to interpret the quantitative data.

keywords: constructional network, productivity, schematization, expansion, social context, frequency

1. Introduction

Until Guardamagna (2016a, 2016b), there have been no in-depth studies of the *secundum* NP ('according to' NP) construction, but only a few fleeting mentions occurring within the framework of grammaticalisation (Ferrari, 1998; Rocha, 1998; Heine & Kuteva, 2003). This preposition has a variety of meanings ranging from space (e.g. 'along') and time (e.g. 'after') to meanings belonging to the domain of logical relationships between participants (e.g. 'in accordance with', 'depending on'). It also expresses two (inter)subjective meanings such as opinion and source of information. A more detailed outlook of these meanings is given in Section 3.

It is presumably because the process of grammaticalisation which led to the development of the preposition *secundum* from the verb *sequor* 'follow' (see Section 2 below) occurred in pre-literary Latin (before BC 285) and cannot there-

https://doi.org/10.1075/cal.21.c7

fore be observed in its step-by-step development that *secundum* NP did not attract much attention within grammaticalisation research. In contrast to the narrow focus of some (early) grammaticalisation studies, mostly centred on the transition from lexis to function word (see Barðdal & Gildea, 2015, p. 2), this paper shows that later attested developments affecting *secundum* NP are also extremely interesting because they highlight the continuous and gradual development underlying constructional change. This paper adopts a constructional and corpus-based approach as a methodological toolkit in order to uncover the changes that *secundum* NP undergoes both in terms of context and frequency. As such, it reflects the evolution, within a construction grammar framework, of some concepts already established within grammaticalisation theory (e.g. "actuation", "host-class expansion", "entrenchment", and "productivity"). In this, my research mirrors the approach taken in Fried (2013), Traugott & Trousdale (2013), and Traugott (2015). This study reveals partial generalisations which "are reflected in usage patterns of individual constructions" and "form a loose but structured network" (Hilpert, 2013b, p. 2). For an overview of the evolution of the concept of grammaticalisation and a discussion of the multiple definitions proposed, see Hopper and Traugott (2003, pp. 19–30), Heine (2004), Heine and Narrog (2011, pp. 1–16), and Traugott & Trousdale (2013, pp. 96–112) among others. Here I rely on Traugott & Trousdale's re-casting of grammaticalisation from a constructional perspective, which has the advantage of unifying two apparently opposing strands in the understanding of the phenomenon, referred to as "grammaticalisation as reduction" (Lehmann, 1995[1982]) and "grammaticalisation as expansion" (Himmelmann, 2004) (see the discussion in Traugott & Trousdale, 2013, pp. 96–112).

Frequency has been a central notion in grammaticalisation studies since at least Bybee's (1985) research highlighting the effects of repetition on language change. The importance of frequency has been acknowledged also within studies in constructionalisation and constructional change, which have developed the field of grammaticalisation from a construction grammar point of view. Under a constructionist approach, changes in form-meaning pairings can be seen as alterations in structure, function and frequency (Hilpert, 2013a, p. 458; see also Hilpert, 2013b, p. 16). A major role is played by frequency data, because "[e]ven in the absence of structural change, a construction may undergo changes in absolute frequency, relative frequency or type frequency. These changes are no less indicative than changes in the substance of a construction" (Hilpert, 2013a, p. 459). At the same time, as Hilpert points out, quantitative analysis does not "obviate the need for close philological analysis" (2013b, p. 6), which may reveal the idiosyncratic development of a construction. In this vein, this paper also delves into the

socio-cultural history of *secundum* NP and its development in relation to Stoic philosophy first and Early Christianity later.

This paper is organised as follows. Section 2 reconstructs the grammaticalisation of *secundum* NP, which is assumed in the literature (Rocha, 1998; Ferrari, 1998; Heine & Kuteva, 2002). Next, Section 3 describes the *Latin Library* corpus and my data sampling while also providing an overview of the semantics of *secundum* NP. Then, Section 4 recalls key concepts related to frequency, productivity and schematicity. Section 5 applies *hapax legomena* counts to the *secundum* NP construction. Section 6 deals with the syntactic and host-class expansion of *secundum* NP and presents networks for conformity sub-constructions. Sections 7 and 8 provide an explanation for the observed host-class expansion. Finally, Section 9 concludes the paper.

2. Grammatical constructionalisation of *secundum* NP: A reconstruction

This section presents a reconstruction of the grammaticalisation/grammatical constructionalisation of the *secundum* NP prepositional phrase which took place in pre-literary Latin.

The preposition *secundum* derives from the gerund[1] (Kühner & Holzweissig, 1966[1912], p. 935, quoted in Heine & Kuteva, 2002, p. 139) or the gerundive[2] (Ferrari, 1998) of the verb *sequor* 'follow'.[3] The available historical data do not show the moment of 'functional split' (Heine & Reh, 1984, pp. 57–59) through which *secundum*, presumably used at some point as both a verb and a preposition, came to be used exclusively as a preposition. Probably, this divergence occurred in pre-literary Latin (that is, before BC 285), when most prepositions are likely

1. The gerund is a verbal noun. It is formed by adding *-(ie)nd-* followed by the endings of the 2nd declension to the present stem of the verb. The gerund has active force, e.g. *modus vivendi* 'a way/style of living'. [The accusative of the gerund ends in *–um.*]

2. The gerundive is a verbal adjective. It is formed by adding *–(ie)nd-us, -a, -um* to the present stem of the verb. It declines like an adjective of the first class (i.e. first/second declension). It is passive in force, e.g. *Carthago delenda est* 'Carthage must be defeated'.

3. As Poultney (1980, p. 34) says, the origins of the gerund and gerundive "have been the subject of extensive debate". This debate focussed on three areas: "the phonology of the characteristic feature of the formation in question, the -nd- […] the interrelated questions of the development of their functions[,] and the priority of origin of one against the other" (ibid.). Also, Vineis (1998, p. 308) points out that there is a debate on whether "the gerund emerged from the gerundive, or whether the entire functional paradigm of the gerundive developed from original forms of the gerund". Therefore the relationship between the gerund *secundum* and the gerundive *secundus, -a, -um is* still quite controversial.

to have developed (Hammond, 1976, p. 132). The form *sequendum*, which in all probability existed as a variant of *secundum* sometime in pre-literary Latin, is a later innovation, as shown by the regular stem vowel *-e-* (*seque-* vs *sequu-/secu-*) modelled by analogy on the present participle[4] (*seque-ns* 'following') (Lindsey, 1915, p. 125; Vineis, 1998, p. 308).[5]

This new form restores the distinction between an active gerund (*sequendum* 'about to follow') and a passive gerundive (*sequendus* 'to be followed'), which the archaic gerundive *secundus* 'about to follow', being anomalously active, did not allow.[6] This is a case of analogical regularisation of paradigms. The change *secundum* (V) > *secundum* (V-Gerund/P) > *secundum* (P) results from a shift in the assignment of the *secundum* NP construction to the relevant schema. Originally, *secundum* NP was presumably licensed by the schema V $NP_{Direct\ Object}$ and later by the P NP schema.

This change in form was accompanied by a change in meaning, from 'follow' to '(coming immediately) after', shown in (1)–(2); 'along' (3); and 'in conformity with' (4) (examples are from the Early Latin sub-corpus of the *Latin Library*, BC 285–149).

(1) *Ite hac* ***secundum vos*** *me.*
Go.PRS.IMP there following you.NOM.PL me.ACC
'You (plural) go there, following me/com- (Plautus, *Stichus*, Act III, Scene I, 45)
ing right after me'

(2) *Ser-ito* ***secundum aequinocti-um*** *vern-um.*
Sow-FUT.IMP after equinox-ACC.N.SG of.spring-ACC.N.SG
'Sow after the spring equinox' (Cato, *De Re Rustica*, 161, 2)

---

4. The present participle is a verbal adjective. It is formed by adding *–(ie)ns/-(ie)ntis* to the present stem. It is declined like an adjective of the second class (3rd declension). Its voice is active, e.g. *mulier clamans discessit* 'the woman left while shouting'.

5. This contrast is also found in the gerundive *agundus,* archaic for *agendus* ('which must be done'), both from ago 'I do'.

6. Gerundives are usually passive in voice, *secundus* being an exception. According to Poultney (1980), the *-nd-* suffix is not a reflex of Indo-European *-tn-* but of *-ny-*, therefore leading to the reconstructed Indo-European form **sekw-on-yos* for Latin *secundus*. The suffix *-(i)yo-* is described as an adjectival suffix, whose value is "not exclusively passive" (ibid., p. 39). For Poultney (1980, pp. 39–40), "the existence in Latin of an active participle probably tended to restrict the use of the formation in *-ndus-* with active sense, leaving it free to develop its passive and obligative sense". The presence of active gerundives like *secundus* or *moribundus* (< infinitive *moriri* 'to die', 'about to die'), though rare, can be explained with reference to the potential for *-ndo-* to develop an active meaning, whose extention to other verbs is 'curbed' by the competing force of the present participle.

(3) *Quid illuc est… homin-um* ***secundum lit-us?***
What thither is man-GEN.M.PL along shore-ACC.N.SG
'Who are those men along the shore?' (Plautus, *Rudens* Act 1, Scene 2)

(4) *Conlaudav-i* ***secundum…*** ***virtut-es***
Praise(PRF)-IND.PRF.1P.SG in.accordance.with virtue-ACC.F.PL
tu-as… (Terentius, *Eunuchus*, Act 5)
POSS.2P.SG-ACC.F.PL
'I praised [you] in accordance with your virtues…'

The emergence of a new conventionalised form-meaning pairing (a new construction) is a case of *constructionalisation*, a process of language change defined by Traugott & Trousdale (2013, p. 22) as "the creation of form$_{new}$-meaning$_{new}$ (combinations of) signs". Constructionalisation minimally involves the "neoanalysis[7] of morphosyntactic form and semantic/pragmatic meaning" (ibid.). Since a preposition is a more procedural unit than a lexical verb, the process of change undergone by *secundum* can be considered a case of *grammatical constructionalisation.* An anonymous reviewer sharply pointed out that *secundum* – historically linked to the gerund(ive) *secundum* – is actually a conservative form, not a new form. My study reflects the position held in the grammaticalisation literature, that the preposition *secundum* arose from the verb *sequor* 'follow' via grammaticalisation, i.e. a process involving both form and meaning (Heine & Kuteva, 2002, p. 139; Ferrari, 1998; Rocha, 1998). The actual dynamic of change from *sequor* (V) to *secundum* (P) is after all still highly speculative.

The growth of a new construction is often paralleled by the obsolescence and loss of a previous construction expressing the same (or similar) meaning, with which it is typically in competition for some time (Traugott & Trousdale, 2013, pp. 71–72). The development of the new preposition *secundum* may have therefore coincided with the disappearance of the construction [[*secundum* Verb]] [follow]],[8] one which is not documented in my data. The database for my investigation is presented in Section 3.

7. The term 'neoanalysis' in Traugott & Trousdale (2013, p. 44) corresponds to 'reanalysis' in the literature on grammaticalisation without any significant difference in meaning.

8. For this notation, see Traugott & Trousdale (2013, p. 8).

3. The *secundum* NP construction in the *Latin Library* corpus

The analysis in this paper is based on data from a 6M sub-section extracted out of the prose component of the 12M POS tagged *Latin Library* corpus (https://cqpweb.lancs.ac.uk/). My study traces the development of *secundum* NP over four periods, thus covering about 800 years: Classical Latin (BC 106 – AD 17), Silver Latin (100–258), Late Latin (330–469) and Early Medieval Latin (530–704).

The analysis presented here is based on a sample of 250 good/clean instances for three out of the four periods examined (Classical, Silver and Late). These instances have been randomly extracted from the total number of occurrences responding to the query *secundum*_PREP in the Latin Library corpus, that is, 315 instances for Classical Latin, 429 for Silver Latin and 1,586 for Late Latin. Of these, 309 were examined in order to get the 250 good/clean instances analysed in my sample for Classical Latin, 271 for Silver Latin, 261 for Late Latin. 'Good/clean instances' means occurrences of *secundum* as a preposition, excluding instances of the string functioning as an adverb ('secondly'), adjective, or pronoun ('the second (one)'), which were erroneously retrieved. Duplicates caused by a glitch in the system have also been discarded. The error rate (number of errors divided by the occurrences examined) was applied to the instances of each period (as a whole) in order to obtain an estimate of the number of good/clean occurrences in each period. Confidence intervals were calculated at the 95% level, thus leading one to estimate that the number of good/clean instances of *secundum* as a preposition is 253–257 for the Classical period, 387–404 for Silver and 1,483–1,555 for Late Latin. The Early Medieval period contained only 161 occurrences, of which only 140 were good/clean. All these 140 instances were considered in this study, which therefore includes the whole population of *secundum* NP occurring in the Early Medieval corpus, and not just a random sample. As a result of the sampling procedure explained above, the analysis in this paper is based on 890 instances of *secundum* NP, which is a representative sample of all the estimated occurrences of *secundum* NP in the corpus considered (estimated to range from 2,263 to 2,356 occurrences).

One methodological issue to be considered here is that using a targeted query for the preposition *secundum* does not allow one to capture the actual usage of the string *secundum* occurring in the various constructions in which it appears (GER. *secundus, ACC. secundum* 'about to follow'; GERV. *secundum* 'following'; ADJ/PRO *secundus* '(the) second' and especially ADV *secundum* 'secondly, after') thus missing out on a full account of the rich phenomena of productivity, competition, and analogy involving *secundum*. This could perhaps be the focus of a larger study: for reasons of space, the present investigation focuses on changes

within the PP only. As a last methodological remark, no ambiguous usages (PREP and GERV. or ADJ/PRO or ADV) were identified in my sample.

A diachronic semantic analysis of the *secundum* NP construction reveals that it underwent major changes over time (see Guardamagna, 2016a, Chapters 4–6; and Guardamagna, 2016b): the most important meanings of *secundum* NP are briefly summarised and exemplified in Tables 1 and 2.

Table 1. Meanings of *secundum* NP and *secundum quod*

Meaning	Example	Translation
Conformity	*secundum naturam vivere* (Cicero, *De Finibus* 5, 24)	To live in accordance with nature.
Limitation	*Quem… secundum carnem… hominem accipimus* (Novatian, *De Trinitate* 17, 3)	(Christ) whom we have received according to his human nature
Space	*Si erunt moenia secundum mare* (Vitruvius, *De Architectura* 1, 7, 1)	If the city walls are along the sea.
Time	*Secundum comitia* … consules … profecti. (Livy, *Ab Urbe Condita* 39, 4)	Close upon the election… the consuls left.
Co-variation	*…donec reddat hominibus secundum actus suos* (Jerome, *Biblia Vulgata*, Ecclesiasticus 35, 24)	Until he has rendered to men according to their deeds/in proportion to their deeds.
Beneficiary	*Secundum te iudicatum erit.* (Aulus Gellius, *Noctes Atticae* 5, 10, 10)	It will be judged in your favour.
Ranking	*Secundum te nihil est mihi amicior solitudine.* (Cicero, *Ad Atticum* 12, 15)	After you/next to you, there is nothing friendlier to me than solitude.
Conjunction	*Quia et nos secundum quod mente aliquid aeternum quantum possumus capimus, non in hoc mundo sumus…* (Augustin, *De Trinitate* 4, 28)	Because we too are not of this world because we reach something eternal with the spirit.

Spatio-temporal meanings are attested and quite frequent in the Classical period but then disappear. Meanings related to the domain of logical relationships between participants are attested throughout. The most token-frequent meanings are the conformity and the limitation ones, whereas the meanings of co-variation, beneficiary, ranking and causation are rare. Finally, the (inter)subjective meanings labelled 'reportative evidential' and 'attribution' are late developments (from Silver Latin onwards) and they are quite rare overall. The (metaphorical and metonymic) relationship between meanings mentioned at the end of Section 3 is discussed in detail in Guardamagna (2016a, Chapters 4–6) and Guardamagna (2016b) where it is also illustrated with a semantic map.

Table 2. Extended conformity meanings of *secundum* NP

Extensions of conformity	Example	Translation
Conformity-Limitation	*Antea enim dictum est ex voce passionis eius, quia secundum carnem cum demonstraret humilitatis exemplum* (Augustin, *Sermones* 75)	Previously in fact (it) was said with the voice of his passion, because in accordance with his flesh/as far as his flesh is concerned when he demonstrated the example of humility.
Conformity-Limitation-Attribution	*Dies secundum Aegyptios inchoat ab occasu solis.* (Isidore, *Etymologiae* 5, 4)	The day according to the Egyptians begins at sunset.
Conformity-Reportative	*Credo… in unum Dominum Iesum Christum… qui…resurrexit tertia die, secundum Scripturas* . (*Symbolum*, *Missale Romanum*, *editio typica tertia*, 2002)	I believe in Jesus Christ, our (lit. the only) Lord, who rose <from the dead> on the third day, according to the Scriptures.
Conformity-Causation	*Alia secundum inmaturam et intemperatam creationem, sicut hi qui dentati nascuntur sive barbati vel cani.* (Isidore, *Etymologiae* 11, 3, 10)	Other [portents] (occur) in accordance with/because of a feature that is premature or untimely, as those who are born with teeth or a beard or white hair.

Table 2 shows meanings derived from conformity and not completely separated from it: in other words these instances can be given both a conformity reading and a reading of limitation, limitation/attribution, reportative evidentiality or causation.

Alongside the spatio-temporal meanings, *secundum* NP expresses logical relations among participants (conformity, co-variation, limitation, conformity-limitation, conformity-causation, beneficiary and ranking). Conformity-limitation-attribution and conformity-reportative afford alternative interpretations depending on context (logical relations and (inter)subjective meanings). *Secundum* NP

is a manner modifier of the VP in spatio-temporal and logical relations, whereas (inter)subjective meanings display sentential scope. The function of conjunction is restricted to the extremely marginal expression *secundum quod* (which is not discussed in this paper).

4. Frequency effects: Productivity and schematicity

The frequency of a construction can be understood in terms of token frequency or type frequency. Token frequency refers to the number of times a given construction appears in a corpus (Bybee, 1985, p. 119, p. 133, 2001, p. 10, 2003, p. 605, 2007a, p. 8, 2007b, p. 269). On the other hand, type frequency "counts how many different lexical items a certain pattern or construction is applicable to" (Bybee, 2007b, p. 269). These two phenomena are responsible for different types of frequency effects. Token frequency is important to determine the unit-status of a construction: high token frequency is taken to correspond to a high level of entrenchment of a construction in the speaker's mind (Langacker, 1987, 1991; Bybee & Thompson, 1997, p. 50; Croft & Cruse, 2004, p. 292; Bybee, 2007a, pp. 13–14; Traugott & Trousdale, 2013, p. 56). Type frequency plays a key role in determining the degree of productivity of a construction and its schematisation.

Barðdal (2008, pp. 19–24) offers an in-depth analysis of the concept of productivity, and argues that the concept can be broken down into lower level notions, namely generality, regularity and extensibility. The concept of generality refers both to the fact that a construction has a wide coverage and is perceived as default, and to the fact that a construction is schematic. The notion of regularity refers to the lack of exceptions to or constraints on a pattern, so that new formations are totally predictable. Finally, extensibility indicates the likelihood that a category extends to both new and nonce items. Among these three concepts, Barðdal argues that "the extensibility concept is of primary theoretical status whereas the generality and the regularity concepts are secondary" (2008, p. 174). The notion of productivity is scalar and its values can be measured (Barðdal, 2008, pp. 24–29; Zeldes, 2012). Typical measurements of productivity involve type frequency.

According to Traugott & Trousdale (2013, p. 114), an increase in productivity and increase in schematicity of a construction, i.e. two of the three factors acting together in grammatical constructionalisation, are compatible with Himmelmann's (2004) concept of expansion, in turn comprising three phenomena: (1) host-class expansion, that is, the increase in an expression's collocational range; (2) syntactic expansion, i.e. the increase in the syntactic uses of an expression; and (3) semantic-pragmatic expansion, in particular the development

of polysemies (or "heterosemies", Lichtenberk, 1991). As Traugott & Trousdale (2013, p. 109) suggest, "all types of expansion may continue after constructionalisation, most especially host-class expansion and syntactic expansion." Indeed, even if my data do not show the constructionalisation process leading to the emergence of the preposition *secundum* (see Section 2), they still show some constructional changes reflecting Himmelmann's expansion (see Section 6). Host-class and syntactic expansion rely on type frequency counts (Zeldes, 2012, p. 98). Productivity, understood as extensibility to new constructions, correlates with high type frequency of a construction. In the case of *secundum* NP, a rise in productivity is measured as an increase in the number of semantic classes to which the NP-arguments belong (host-class expansion) and an increase in the range of constituents that *secundum* NP modifies (syntactic expansion). As pointed out by Traugott & Trousdale (2013, p. 114), expansion is a sign of a construction's productivity and increased schematicity.

Beyond type frequency counts, other measures for productivity have been elaborated in the literature (see Barðdal, 2012, Chapter 2 and Zeldes, 2012, Chapters 3–4 for a complete overview). Not all of these measures can be applied to *secundum* NP.[9] My observations are based on *hapax legomena*, and specifically measured using the *category conditioned degree of productivity* index, also known as Baayen's P (Baayen, 2009, p. 902; Zeldes, 2012, p. 60). Baayen's P is a widely used productivity measure in morphology, which "has also been applied in some of the few studies addressing syntactic issues (Kiss, 2007; Zeldes, 2009, 2011)" (Zeldes, 2012, p. 64). This measure is obtained by dividing the total number of *hapax legomena*, i.e. constructions with a token frequency equal to 1, by the total number of tokens of the category in question (in my case the *secundum NP* construction). This index expresses the likelihood of a process "to produce new members, or how saturated a process is" (ibid., p. 63). In other words, it "qualifies the potential for innovation, based on the probability of a new previously unforeseen type, as opposed to the repetition of familiar types" (ibid., p. 68). As pointed out by Barðdal (2008, p. 27), "hapaxes reflect the productivity (in the sense of 'combinability')" of a particular morphological or syntactic pattern, whereas lexical items with a high token frequency are more likely stored in the memory as non-analysed

9. For instance, the *hapax conditioned degree of productivity* (P*, Baayen, 1993) is not easily applicable to my case study. In morphology, P* is calculated as "the number of hapaxes formed with a certain affix in a corpus divided by the global number of hapaxes within the corpus" (Gaeta & Ricca, 2006, p. 61). These figures are not easily retrievable within the *Latin Library* corpus.

units. *Hapax legomena*, generated on the fly[10] are considered to be reliable indicators that a pattern is productive. Baayen's P is always a number between 0 and 1, the 0 meaning that there are no *hapax legomena* at all, and 1 meaning that all the instances of the category are *hapax legomena*. Therefore, the higher Baayen's P, the more productive the construction.

Alongside type frequency counts, Barðdal argues that a key role in measuring productivity is played by coherence, i.e. "the internal consistency found between relevant items" (2008, p. 172), be it phonetic or semantic. Therefore, Barðdal (2012, p. 469) defines productivity as the result of both type frequency and coherence:

> For a category that is high in type frequency, this means that only a low degree of internal consistency across its items is needed for the category to be extended. Categories of intermediate type frequency have to be more coherent than high type frequency categories in order to be extended to new items. Finally, categories that are low in type frequency must show the highest degree of internal consistency across their items in order to be extended to new types.

On the basis of this, productivity can be thought of as a cline, as represented in Figure 1.

The vertical axis represents type frequency and the horizontal one represents coherence. The oblique line shows that high productivity correlates (a, top left) with items that have high type frequency even if they have low semantic coherence, and (b, bottom right) with items that have low type frequency but display a high degree of semantic/phonetic coherence. Analogy plays a key role in the spread of type (b). Since productivity are represented as a cline, there are intermediate degrees of productivity, correlating with intermediate levels of type frequency and coherence.

---

10. Baayen (2009, p. 906), talking about morphology, but the same principle may be applied to syntax as well, observes that "words with less compositional meanings typically tend to be high frequency words." In the same vein, Zeldes (2012, p. 73) claims that *hapax legomena* are "produced compositionally".

As pointed out by Joel Olofsson (personal communication) the role of compositionality in Construction Grammar is a debated issue. Specifically, Goldberg (1995) shows that entire structures (e.g. the ditransitive construction, the way-construction) can also contribute meaning beyond the meaning of the lexical constituents that fill its slots.

Kay & Michaelis (2012) is a reflection on the topic of compositionality in Construction Grammar, through which they address "the incorrect supposition" that "constructional approaches are opposed to compositional semantics" (p. 2271). However, one shortcoming of this paper is that the authors do not distinguish between type and token frequency: when they state that there is a continuum/gradient from idiomaticity to productivity they refer to type frequency, whereas where they observe that also idiomatic constructions (like the pseudo-conditional they discuss) can be quite productive they refer to token frequency.

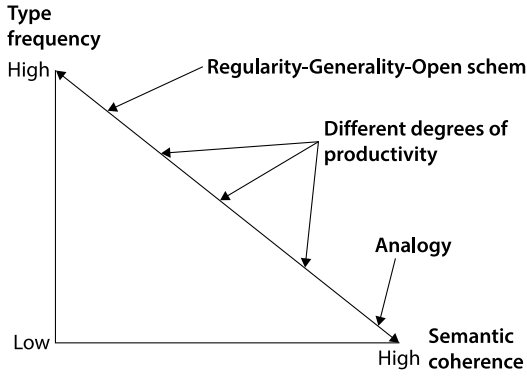


Figure 1. The inverse relationship between coherence and type frequency (Barðdal, 2008, p. 38)

Barðdal's (2008, p. 38) model contributes toward explaining the increased productivity of *secundum* NP in the transition between the Classical and Silver periods (see Section 7).

5. *Hapax legomena* counts and productivity

Table 3 presents the number of *hapax legomena*, types, tokens and Baayen's P (*hapax legomena*/tokens) value for each period considered independently.

Table 3. Productivity of the *secundum* NP construction (including *secundum quod*) in the *Latin Library* corpus

***Secundum* NP**	**Classical**	**Silver**	**Late**	**Early Medieval**
Hapax legomena	70	92	72	57
Total tokens	250	250	250	140
Total types	104	137	120	84
Baayen's P (*hapax legomena*/tokens)	0.28	0.37	0.29	0.41

In order to interpret the results in Table 3, a series of critical preliminary observations are in order. First, it can be noticed that for each period the value of Baayen's P is very high – actually higher by a factor of 10 compared to studies in morphology such as Gaeta & Ricca (2006), Štichauer (2009) and many studies discussed in Zeldes (2012). However, it seems that this measure is not implausible when measuring syntactic productivity as shown for instance in Zeldes's study of the German adposition *wegen* (2012, p. 109), which reaches similar values to the

ones displayed by *secundum*. This may perhaps suggest that Baayen's P indices could vary depending on whether one deals with morphology or syntax, an intuition which should of course be borne out by further research.

The Baayen's P index, examined period by period, seems to suggest that *secundum* NP is most productive in the Early Medieval period (530–704), whereas Late Latin (330–469) shows a level of productivity similar to the Classical era (BC 106 BC – AD 17). Moreover, Silver Latin (100–258) shows an increase compared to the Classical period. Still, conclusions based on this data may be premature because they do not take into account the distorting effect of different sample sizes on Baayen's P. In examining a corpus (or a sample), new types are easily found as the number of tokens grows, but this growth slows as the sample size increases since in a larger sample it is more likely that new tokens are merely occurrences of already attested types (Štichauer, 2009, p. 139, quoting Baayen, 2008). Comparing different sample sizes therefore has a relevant effect on *hapax legomena* counts and – as a consequence – also on Baayen's P. For this reason, the figures for the Early Medieval period could be "inflated" because of the smaller sample size as pointed out in Gaeta and Ricca (2006, pp. 58–59).

As mentioned in Section 3, I examined 250 random tokens[11] for Classical, Silver and Late Latin[12] and only 140 instances for the Early Medieval period (i.e. the whole population for that period). Providing data for a larger study on the preposition *secundum* not specifically aimed at productivity (Guardamagna, 2016a, 2016b), the four sub-corpora (Classical, Silver, Late and Early Medieval) have been designed in such a way as to have a similar temporal extension but not a similar word count, and this can have an impact on *hapax legomena* and type counts. All these caveats suggest that the observations based on Baayen's P need to be interpreted with caution.

The Baayen's P index raises some further issues. First of all, it is not necessarily easy to determine a threshold for what counts as a high or low proportion. Specifically, it is unclear what the threshold levels of Baayen's P are for determining some degree of productivity, high or low. A similar problem has been pointed out in

11. In this paper, I use the term 'sample' to refer to the number of tokens of *secundum* NP I analysed out of all the tokens of *secundum* NP present in a sub-corpus. This use differs from the meaning of the term 'sample' in works on productivity, where it identifies similar-sized sections of a corpus which are analysed in their entirety.

12. These 250 token samples differ in how representative they are with respect to the estimated total number of instances of the preposition *secundum* in the four sub-corpora. As stated in Section 3, the projected actual matches in the sub-corpora are 255 (confidence interval 253–257) for the Classical period, 396 (confidence interval 387–404) for Silver and 1,519 (confidence interval 1,483–1,555) for Late Latin; all the instances have been considered for the Early Medieval period, i.e. the whole population is considered, not just a sample.

Bybee (2007a, p. 16) regarding the operationalisation of the concepts of high/low type and token frequencies, and by Hollmann (2003, p. 110) specifically in relation to productivity. Furthermore, when comparing Baayen's P results (across periods or between constructions), it is not clear how to interpret the variation between indices. In particular, when the observed differences are small it may not be easy to assess their impact on productivity. Another complication relates to the fact that ordinary significance tests are not appropriate for determining whether the variation captured by *hapax legomena* indices is significant.

Additionally, variation in the indices could be linked to a different genre distribution within the various sub-corpora. For instance, the dip in Late Latin is likely to be related to the composition of that period's sub-corpus in which the *Bible* and the *Codex Theodosianus* distort the data by reflecting a conservative use of the construction. The lower productivity of *secundum* NP in the Classical period may be explained by the idiomatic use of the token-frequent conformity construction *secundum naturam*, accounting for 27% of the data (68/250 instances). I return to this construction in Section 8. Regarding the interpretation of the increase of types in the Silver period, some insights may come from the analysis provided in Section 6.2.

For all these reasons, the observations based on Baayen's P outlined so far should interpreted with care.

More promising results than the period-by-period analysis can be reached through considering the corpus as a whole, instead of the single sub-corpora. Table 4 shows the progressive merging of *hapax legomena* and types, obtained by considering first the Classical period, then Classical and Silver together, then the Classical, Silver and Late periods combined, and finally all four periods as a whole. This way, both *hapax legomena* and other types (with more tokens) which are already attested in a previous period disappear from the count – with an effect of reduction with respect to the mere sum of types and *hapax legomena* for the periods considered.

Table 4. Progressive calculation of *hapax legomena*, types, and Baayen's P for *secundum* NP in the *Latin Library* corpus

***Secundum* NP**	**Classical**	**Classical & Silver**	**Classical, Silver & Late**	**Classical, Silver, Late & Early Medieval**
Hapax legomena	70	150	179	204
Types	104	222	293	335
Tokens	250	500	750	890
Baayen's P	0.28	0.30	0.24	0.23

On the basis of the data in Table 4, it is possible to plot the vocabulary growth curve for *secundum* NP. The graph in Figure 2 represents the growth of both the *hapax legomena* (HL) and the types of *secundum* NP as a function of the number of tokens N. The graph almost corresponds to a monotonically increasing function, showing a sharper increase and then slowing down as the sample size increases. However, the values for the Classical and Silver period (merged) appear to be slightly higher than expected. This reflects the aforementioned anomaly of the Silver period.

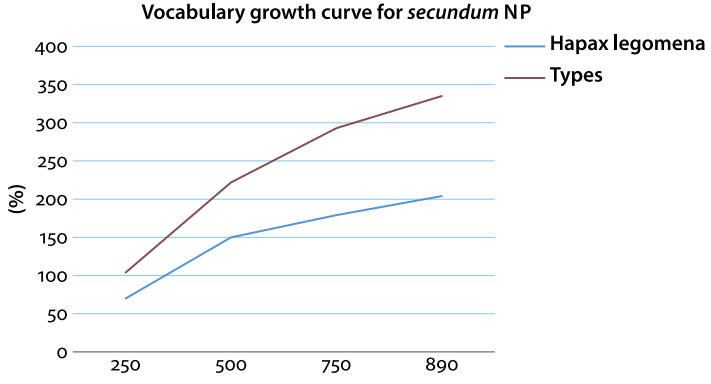


Figure 2. Vocabulary growth curve for *secundum* NP in the *Latin Library* corpus. The x axis represents the number of tokens obtained by adding together the sub-corpora samples. The y axis represents the number of types/*hapax legomena*

The vocabulary growth curve for the construction seems to suggest that this productivity of *secundum* NP remains high throughout the period considered, which is compatible with claims about post-constructionalisation constructional changes. In order to yield more reliable results, a further study should be based on a corpus controlled for sample size (total words per period), and possibly balanced and representative in terms of authors and genres.

6. Type frequency counts: Syntactic and host-class expansion of *secundum* NP

In what follows I concentrate on *secundum* NP's conformity constructions, and consider its host-class and syntactic expansion. I focus on this meaning because it is represented across all four periods and its token frequency is high enough across the whole corpus to allow us to observe clear patterns.

6.1 Syntactic expansion

Secundum NP as an adverbial is a modifier scoping over constituents below the sentential level or coinciding with the sentence. In the examples below, *secundum* NP scopes over constituents below the sentence, as in (5) PP *in capite* 'in the head' and (6) V *fiunt* 'occur', and over the sentence itself (7). In sentence (8), *secundum quod* can be interpreted as a conjunction.

(5) *Vuln-us* *accep-it* *in capit-e* ***secundum***
Wound-NOM.N.SG receive(PRF)-PRF.3P.SG in head-ABL.NSG behind
aur-em (Sulpic. ad Ciceronem)
ear-ACC.F.SG
'He received a wound in the head behind his ear'

(6) *Omn-ia* *qu-ae* ***secundum natur-am***
All-ACC.N.PL REL-NOM.F.SG according.to nature-ACC.F.SG
fi-unt *sunt* *hab-end-a* *in bon-is*
happen-IND.PRS.3P.SG be.IND.PRS.3P.SG have-GERV-NOM.N.PL in good-ABL.N.PL
'Everything that happens in accordance with (Cicero, *De Senectute* XIX, 71)
nature is to be reckoned among blessings'

(7) *Quia post resurrection-em* *eius* *a* *mortu-is*
Why after resurrection-ACC.F.SG he.GEN.3P.SG from dead-ABL.M.PL
qu-ae *die* *terti-a* *effect-a* *est cael-i*
REL-NOM.F.SG day(ABL.F.SG) third-ABL.F.SG done-NOM.F.SG is heaven-NOM.M.PL
eum *recep-erunt* ***secundum*** ***propheti-am?***
he.ACC.3P.SG receive(PRF)-PRF.3P.PL according.to prophecy-ACC.F.SG
'Why after the resurrection from the dead (Tertullian, *Adversus Iudaeos*, 13, 23)
did the heavens receive him back, according to the prophecy/as the prophecy
says?'

(8) *Quia* *et* *nos* ***secundum*** ***quod*** *ment-e*
Because also we.NOM in.accordance.with REL.ACC.N.SG mind-ABL.F.SG
aliquid *aetern-um…* *capi-mus,* *non in hoc*
something.ACC.N eternal-ACC.N.SG reach-IND.PRS.1P.PL not in this.ABL.M.SG
mund-o *sumus…*
world-ABL.M.SG be.IND.PRS.1P.PL
'Because we too are (Augustin, *De Trinitate*, 4, 28, translated by Hadden, 1873)
not of this world because we reach something eternal with the spirit'

As Table 5 shows, *secundum* NP is almost exclusively a predicate modifier in Classical Latin, but in Silver Latin it expands its scope to a variety of constituents below the clause (AP, NP, PP) and to the clause itself (S). In addition to *secundum* NP

constructions scoping exclusively on the VP or the S, in Silver and Late Latin there are also contexts in which the scope can be interpreted either as over the predicate or over the sentence (labelled as VP/S in Table 5).

Table 5. Scope of *secundum* NP in the *Latin Library* sample

Scope	Classical	%	Silver	%	Late	%	Early Medieval	Early Medieval (normed)	%
AP	0	0	6	2.4	1	0.4	0	0	0.0
NP	44	17.6	21	8.4	5	2	3	5	2.0
PP	0	0	4	1.6	0	0.0	0	0	0.0
VP	206	82.4	165	66	228	91.2	111	198	79.2
VP/S	0	0	23	9.2	2	0.8	0	0	0.0
S	0	0	30	12.0	13	5.2	26	47	18.8
Conjunction	0	0	1	0.4	1	0.4	0	0	0.0
Total	**250**	100	**250**	100.0	**250**	100.0	**140**	**250**	100.0

Even if a *chi-squared* significance test suggests that the changes in scope in Table 5 are very highly significant[13] ($\chi^2 = 208.824$, $df = 18$, $p < 0.00001$), applying the test is problematic because of the many cells in which the expected frequencies are below one or equal to zero. Excluding Classical Latin does not improve the situation, so that the data remain problematic for the application of the significance test. In order to increase the reliability of *chi-square* testing, the *chi-square* has been applied only to the NP, VP and S values (consolidated in Table 6). The result ($\chi^2 = 113.798$, $df = 6$, $p < 0.0001$) suggests that the change is extremely significant.

Table 6. Scope of *secundum* NP in the *Latin Library* sample (consolidated table)

Scope	Classical	Silver	Late	Early Medieval	Early Medieval (normed)
NP	44	21	5	3	5
VP	206	165	228	111	198
S	0	30	13	26	47
Total	**250**	**216**	**246**	**140**	**250**

13. In what follows I give absolute frequencies and percentages, and the results of the significance test. I use the *chi-square* test, following the conventions in referring to difference levels indicated as "significant" ($p \leq 0.05$), "highly significant" ($p \leq 0.01$) and "very highly significant" ($p \leq 0.001$) (Butler, 1985).

Table 6 outlines two lines of development which need to be considered in parallel, namely a slight contraction of the instances of *secundum* NP with predicate scope and an increase in the distribution of *secundum* NP with sentential scope.

The high values of VP scope in the Late Latin period in Figure 3, and the dip in the S-scope in this period in Figure 4, may be the result of a distortion caused by the *Codex Theodosianus* and the *Bible*, which lend the Late Latin corpus a 'conservative' character, causing it to resemble the Classical period, as already mentioned in Section 5.

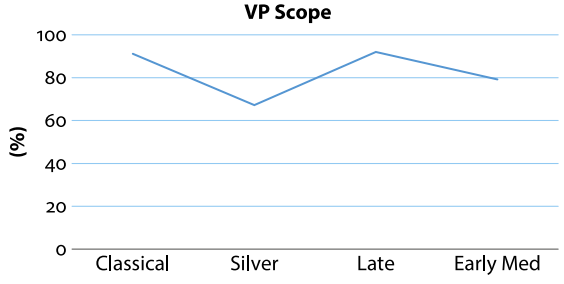


Figure 3. Percentage of *secundum* NP with VP scope over the total sample of *secundum* NP in the *Latin Library* shown by period

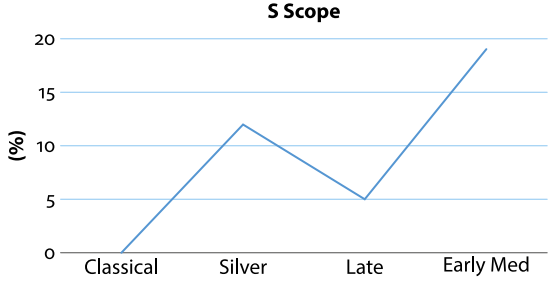


Figure 4. Percentage of *secundum* NP with S scope over the total sample of *secundum* NP in the *Latin Library* shown by period

These data are consistent with several studies on the unidirectionality of the development of adverbs, which demonstrate that manner adverbs tend to precede the appearance of sentential adverbs (e.g. Traugott, 1995; Traugott & Dasher, 2002, p. 40). The increase in scope of an adverb is often linked to its (inter)subjectification: sentence adverbials tend to be more speaker-oriented than adverbials with scope below the sentence, which tend to be more referent-oriented. Among the *secundum* NP constructions expressing extensions of conformity there are expressions developing reportative (Examples 9–10) and attribution meanings (11), which are (inter)subjective in nature.

(9) ***Secundum psalm-um***, *sed-e* *ad dexter-am*
According.to psalm-ACC.M.SG sit-IMP.PRS.2P.SG at right-ACC.F.SG
me-am. (Tertullian, *Adversus Marcionem* 4, 17, 12)
POSS.1P.SG-ACC.F.SG
'According to [what is written in] the psalm: "Sit at my right…"'

(10) *Effund-am de spirit-u me-o in*
Pour.out-IND.FUT.1P.SG down.from spirit-ABL.M.SG POSS.1P.SG-ABL.3P.SG in
omn-em carn-em, ***secundum Ioel-em***.
all-ACC.F.SG flesh-ACC.F.SG according.to Joel-ACC.M.SG
'I will pour out my spirit upon all flesh,(Tertullian, *Adversus Marcionem* 5, 4, 2)
as Joel has it'

(11) *Plebisscit-um igitur est* ***secundum*** *eum* ***Capiton-em***
Plebisscitum-ACC.N.SG therefore is according.to he-ACC.3P.SG Capito-ACC.M.SG
lex, qu-am pleb-es, non popul-us,
law(NOM.F.SG) REL-ACC.F.SG plebs-NOM.F.SG not people-NOM.F.SG
accip-it. (Aulus Gellius, *Noctes Atticae*, 10, 20, 6)
adopt-IND.PRS.3P.SG
'Therefore, according to (this) Capito, a plebisscitum (a decree binding on the plebs) is a law which the commons adopt, and not the people'

To conclude, the data discussed in this section show an increase in the scope of *secundum* NP, which is consistent with its increase in (inter)subjectivity over time (see semantic-pragmatic expansion in Himmelmann, 2004; also see Section 4). The evolution of the *secundum* NP construction supports the observation by Traugott & Trousdale (2013, p. 109) that syntactic expansion is a process of constructional change which may extend well beyond the phase of constructionalisation.

Having discussed syntactic expansion, one of the measures of expansion (and thus extensibility) discussed by Himmelmann (2004) and incorporated in Traugott & Trousdale's (2013) model of constructional change, this paper now turns to host-class expansion.

6.2 Host-class expansion

This section takes a close look at the arguments of *secundum*, with particular attention to their semantic class. My approach is similar to the one found in studies on the grammaticalisation of English *BE going to*, which initially collocated only with dynamic verbs and later with both dynamic and stative verbs (Himmelmann, 2004, p. 33). This section focuses on the transition between the Classical and Silver periods.

Table 7 is organised by dividing the arguments of *secundum* into six semantic[14] classes (nature, law, product of speech and writing, people and humanised God, concrete terms, and abstract terms). The nature class is peculiar, because it has only one member (the noun *naturam* 'nature'): its autonomy from the neighbouring class of law-like terms is discussed below. Examples of arguments taken from the class indicating products of speech and writing are *Scripturas* 'Scriptures', *verba* 'words', and *sententiam* 'statement'. Examples of nouns indicating God are *Deum* 'God' and *creatorem* 'the creator', and terms indicating people are *apostolum* 'apostle' and *Platonem* 'Plato'. Within the category indicated as concrete are nouns such as *solem* 'sun', *carnem* 'flesh', and *arbusta* 'bushes'; and finally, the large category of abstract nouns embraces a variety of items like *fidem* 'faith' and *veritatem* 'truth'.

Table 7. Semantic classes of the arguments of *secundum* (and the syntactic class of pronouns)

Synopsis of argument classes (head noun by lemma)	Classical Latin			Silver Latin		
	Types	Tokens	% token on total conformity tokens	Types	Tokens	% token on total conformity tokens
Nature	1	68	75	1	4	2
Law	9	20	22	9	22	14
Products of speech and writing	0	0	0	9	15	10
God and people	0	0	0	11	17	11
Concrete	0	0	0	9	11	7
Abstract	2	2	2	50	70	45
Pronouns & dummy pronouns	1	1	1	9	17	11
Total instances of conformity	**13**	**91**	100	**98**	**156**	100

Secundum NP in Classical Latin draws mainly from two classes, the nature class and the more varied class of law-like terms (e.g. *tabulas* 'laws', *decretum*

14. The class of pronouns is different from the others, because it is not a semantic class, but a syntactic one. Incidentally, it is worth pointing out that the increase in different types of pronouns indicates that *secundum* syntactically expanded the types of arguments it takes, moving from NPs only to both NPs and PROs.

'decree', *ius* 'law'). These are illustrated in Examples (12) and (13) drawn from a philosophical treatise and a historical narration respectively.

(12) *Ita fin-is bon-orum exist-it,* ***secundum***
Thus end.NOM.F.SG good-GEN.N.SG exist-IND.PRS.3P.SG according.to
natur-am *viv-ere.* (Cicero, *De Finibus* 5, 24)
nature-ACC.F.SG live-INF.PRS
'Thus arises "the end of goods", namely to live in accordance with nature'

(13) *Duumvir-os…* ***secundum leg-em*** *faci-o.*
Duumvir-ACC.M.PL according.to law-ACC.FG.SG do-IND.PRS.1P.SG
'I elect the duumviri (magistrates) according to (Livy, *Ab Urbe Condita* 1, 26)
the law'

There is only one example of the pronoun *hoc* 'this', which refers to a law-like item, and two abstract terms, which trigger a law-like interpretation. In Example (14) *rationem* 'reasoning' refers to a linguistic regularity/rule explained above, and similarly in Example (15) *analogias* 'analogy' (lit. plural 'analogies') refers to a linguistic rule structuring grammar.

(14) *Item* ***secundum*** *ill-orum* ***ratio-nem*** *debe-mus*
Likewise according.to POSS.3P.PL-GEN.M.PL reasoning-ACC.F.SG must-PRS.1P.PL
second-is syllab-is long-is Hector-em
second-ABL.F.PL syllable-ABL.F.PL long-ABL.F.PL Hector-ACC.M.SG
Nestor-em. (Varro, *De lingua Latina* 8, 72)
Nestor-ACC.M.SG
'Likewise, according to their theory, it is with long second syllables that we should pronounce *Hectorem* and *Nestorem*'

(15) *In his verb-is qu-ae contrari-a non*
In this.ABL.N.PL word-ABL.N.PL REL-NOM.N.PL opposite-NOM.N.PL not
hab-ent, ut loqu-or *et* ven-or, *tamen*
have-IND.PRS.3P.PL such.as speak-PRS.1P.DEP and hunt-PRS.1P.DEP nevertheless
dic-imus loqu-ens *et* ven-ans, loc-uturus *et*
say-PRS.1P.PL speak-PRS.PTCP.SG and hunt-PRS.PTCP.SG speak-FUT.PTCP.SG and
ven-atur-us, loc-ut-us *et*
hunt-PTCP.FUT-NOM.M.SG speak-PTCP.PST-NOM.M.SG and
ven-at-us, *quod* ***secundum*** ***analogi-as***
hunt-PTCP.PST-NOM.M.SG REL.NOM.N.SG in.accordance.with analogy-ACC.F.PL
non est, quoniam dic-imus loqu-or *et*
not is because say-IND.PRS.1P.PL speak-IND.PRS.1P.SG.DEP and
ven-or, *non* loqu-o *et* ven-o.
hunt-PRS.1P.SG.DEP not speak-PRS.1P.SG.ACT and hunt-PRS.1P.SG.ACT

> 'In these words, which do not have an opposite (Varro, *De Lingua Latina* 8, 59)
> (voice) like *loquor et venor*, we nevertheless say *loquens* and *venans*, *locuturus*
> and *venaturus*, *locutus* and *venatus*, which does not hold by analogy, because
> we say *loquor* and *venor*, not *loquo* and *veno*'

This confirms the existence of a dominant pattern, namely that the complement of *secundum* with conformity meaning is a noun with a normative semantic component. As explained in detail in Section 8, the noun *naturam* 'nature.ACC' also expresses a normative meaning, since it is to be interpreted as 'the laws of nature'. However, I argue that the *secundum naturam* construction in the Classical period is likely to be perceived by the speaker as an autonomous construction instead of the instantiation of a more general pattern *secundum* + NP_{Law}. This is because of its high token frequency, idiomatic meaning, and tendency to appear with larger fixed constructions in which it modifies a restricted number of predicates. Specifically, *secundum naturam* modifies *vivere* 'to live' (10 instances), *esse* 'to be' (51 instances), and a handful of other verbs (7 occurrences). These verbs are: *dividitur* 'is divided', *divisus est* 'was divided', and *affectum esse* 'is affected' (all in the passive voice); *ad…optinenda* 'to be obtained' (passive obligative meaning); and *fiunt* 'they become' (passive of *facio* 'do, make', lit. 'they are made'). There is only one active verb: *nominant* 'they nominate.' In view of this behaviour, I consider the *secundum naturam* construction in Classical Latin as autonomous and propose the constructional network for the *secundum* NP conformity construction represented in Figure 5.

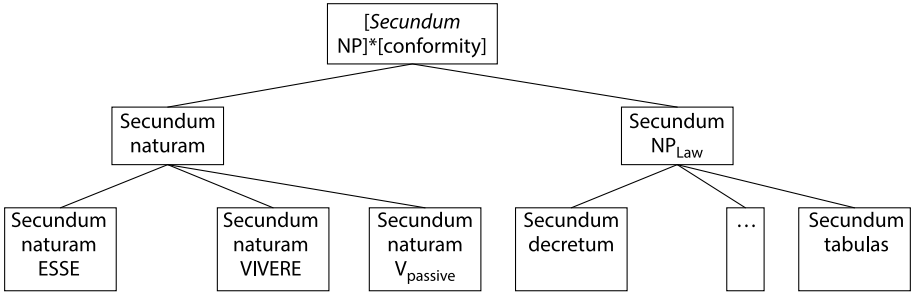


Figure 5. Conformity constructions in Classical Latin: the position of *secundum naturam*

The constructional network for conformity in the Silver Latin period is different (see Figure 6). In fact, *secundum naturam* drops sharply in token frequency (from 68/91 instances down to 4/156, i.e. from 74.7% to 2.6%), thus suggesting perhaps the possibility to consider the construction *secundum naturam* as an instantiation of the more general pattern *secundum* NP_{Law}. The loss of autonomy[15]

15. The idea that a specific construction becomes a part of a more abstract one when it loses token frequency needs to be explored further in Construction Grammar.

of *secundum naturam* could perhaps be considered the effect of the text composition of the Classical Latin sub-corpus in contrast to the Silver Latin one, in turn reflecting a cultural shift from paganism and Stoic philosophy to Christianity (see Section 8).

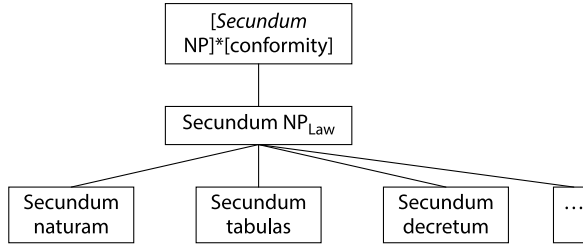


Figure 6. Conformity constructions in Silver/Late/Early Medieval Latin, including the changed status of *secundum naturam*

Another phenomenon affecting the *secundum* NP constructional network in Silver Latin is the expansion of its collocational range, drawing on three further classes not represented in the Classical period: products of speech and writing, people and personified God, and concrete and abstract nouns. Coussé (forthcoming) proposes an interpretation of host-class expansion in semi-schematic constructions in terms of prototype theory. According to this, host-class expansion starts with a core, and moves to include more peripheral members of a category. Based on this suggestion, it can be proposed here that the core elements in the conformity construction are represented by nouns of law/nature. The new members, referred to as products of speech and writing, share the semantic element 'text' connected with the practice of writing a code of laws. Then, the extension to God/people could perhaps be explained as encoding the author of products of speech or writing. Finally, abstract and – even more so – concrete Ns are the most semantically distant from the semantic core represented by nature/law, showing only a very tenuous connection.

The full constructional network for *secundum* NP expressing conformity in Silver Latin is therefore the one represented in Figure 7.

Traugott & Trousdale (2013, p. 28) point out that according to Hilpert (2012, p. 234), the increased collocational range of a construction is a clue to its semantic extension, and in particular its schematisation.

My analysis now turns toward addressing some factors favouring the increase in the productivity of *secundum* NP.

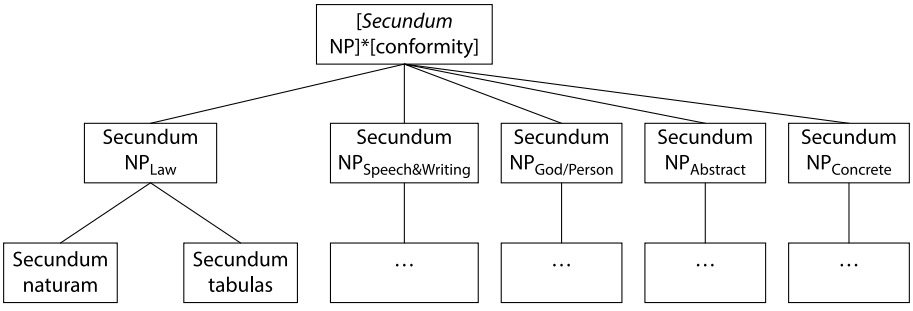


Figure 7. Conformity constructions in Silver and Late Latin. In Early Medieval Latin the category of concrete nouns is not represented

7. Accounting for the increase in productivity with Baŕdal (2008)

As argued in Section 6, type frequency counts record a process of expansion for *secundum* NP. As far as host-class expansion is concerned, type counts show that in the transition between Classical and Silver Latin there is an increase from two to four semantic classes. Regarding syntactic expansion, similarly, there is a transition from one to five types of constituents over which *secundum* NP scopes.

All these measures capture an increase, but they do not point to any possible factors which may explain the cause for such an increase in productivity. To the best of my knowledge, the only model which affords a possible explanation for the observed increase is Barðdal's (2008) model of productivity.

Going back to this model, Barðdal (2008, 2012; also see Section 4) argues that productivity is the result of the inverse relation between type frequency and coherence. In other words, low type frequency constructions are extensible provided they are very coherent, for instance, if they form a semantically coherent group. Under this view, the fact that in Classical Latin *secundum* takes only *naturam* 'nature' or law-like NPs as its argument demonstrates its semantic coherence and explains its productivity, resulting eventually in its expansion to other semantic classes.

As regards *secundum naturam*, a possible explanation of its unexpected productivity lies in the relationship between productivity and analogy. For Barðdal (2008, p. 44), very specific constructions can attract new members by analogy, provided that they are very entrenched. This is the case for *secundum naturam*, with its 68/91 occurrences, accounting for three quarters (74.7%) of all instances of conformity and almost one third (27%) of all instances of *secundum* NP. Analogy, in Barðdal's view, is just "the other side of the productivity coin" (2008, p. 44), i.e. "a type of extension based on only one model form" (ibid.). In Section 8, I discuss specific constructions within the *secundum* NP conformity network and

propose that some other constructions may have been created by analogy with *secundum naturam*.

To sum up, my data show that the *secundum* NP conformity construction expands over time both in syntactic terms and in terms of the semantics of its arguments. In accounting for this expansion, Barðdal's (2008) productivity cline enables one to account for the expansion of the *secundum* NP conformity construction better than models exclusively based on type frequency counts. The following section takes a closer look at *secundum naturam* and some constructions possibly related to it by analogy.

8. Analogy and cultural factors at play

Firstly, the analysis presented in this section addresses the question as to why the *secundum naturam* 'in accordance with nature' construction occupies such a prominent role in Classical Latin, as evidenced by collocational indices. *Naturam* 'nature' is the first collocate to the right of *secundum* in Classical Latin (log-likelihood = 874.897), showing a stronger relation than the second collocate *mare* 'sea' (log-likelihood = 77.491), as suggested by the fact that the difference in their log-likelihood is very highly significant ($p < 0.00001$).

A possible answer for the importance of the *secundum naturam* construction can be found in the specific cultural context in which it was produced, in particular Stoic philosophy. Then, my analysis moves on to consider how this construction may have provided a template for some early Christian expressions, such as *secundum carnem* 'in accordance with the flesh' and *secundum spiritum* 'in accordance with the spirit'.

The construction *secundum naturam* mainly has a philosophical meaning, particularly relevant in Cicero's writings. The construction *secundum naturam vivere* is paraphrased as *convenienter congruenterque naturae vivere* 'to live in agreement and in harmony with nature' (Cicero, *De Finibus* 3, 9, 30; see also 4, 11 and Cicero, *De Officiis* 4, 13), thus indicating consistency (or 'conformity'). The expression *secundum naturam (vivere)* is a calque from Greek *homologouménōs tẽi fýsei* (*zẽn*) 'to live in accordance with nature', a formula introduced by the Hellenistic philosopher Zeno of Elea as 'to live consistently' and later modified as 'consistently with nature'[16] (Vernon Arnold, 2014[1911], p. 282, footnote 68) and as such entered into Stoicism. As Vernon Arnold points out, the expression '"life according to nature' means to the Stoics 'life in accordance with

16. 'To nature' was added by Zeno himself, or Cleanthes or Chrysippus (Vernon Arnold, 2014[1911], p. 282, footnote 68).

the general movement of the universe', to which the particular strivings of the individual must be subordinated" (2014[1911], p. 283). The Good is identified with what is in accordance with nature, so things are classified depending on whether they are according to nature or against nature. Within certain strands of Stoicisim, it is possible to notice that there is a 'normative' element to nature, as shown in Chrysippus's words: "to live virtuously is to live according to the scientific knowledge of the phenomena of Nature, doing nothing that the Universal Law forbids" (Diogenes Laërtius, *Vitae Philosophorum* VIII, 87–88; in Vernon Arnold, 2014[1911], p. 283, footnote 70; see also Cicero, *De Finibus* 3, 9, 30). The link goes both ways, as it is suggested that nature is the source of the law (Cicero, *De Officiis* 3, 72).[17] According to Cicero, the "law of Nature is the single system of law governing the entire universe to which all natural things have an obligation to conform" (Sharma & Karma, 2003, p. 211, see also p. 193). In this view, the construction *secundum tabulas* 'in accordance with the laws' also assumes a more complex meaning: within Stoic philosophy it refers to a political duty, in turn based on the moral duty to obey the laws of nature on which the civil laws are expected to be based. This is particularly relevant in Cicero's works, which present a "synthesis of ethics and politics" (Sharma & Karma, 2003, p. 211).

Going back to the construction *secundum naturam*, what is not 'according to nature' is 'against nature' (see also Cicero, *De Finibus*, 2, 34; 3, 9; 3, 16; 3, 20; 3, 21; 3, 59; 5, 30), as much as what is not 'in accordance with the law' is 'against the law'. As De Boer (2005, p. 209) points out, "in Stoic monistic view, there is no cosmic power contrary to Nature", so things can be either in accordance with nature or contrary to it, but they cannot be in accordance with something which is essentially opposed to nature.

Stoic philosophy strongly influences Early Christian thought (Colish, 1990, p. 9). This heritage is also expressed through language: it is not surprising that the construction *secundum naturam* transitions into Christian thought. For instance, it is found in Ambrose's definition of *decorum* ('decency'), as in (16):

(16) *decor-um* *est* ***secundum*** ***natur-am*** *vivere… et*
Decency-NOM.N.SG is in.accordance.with nature-ACC.F.SG live-INF and
turpe *est* *quod* *sit* *contra* *natur-am.*
vile-NOM.N.SG is what.ACC.N.SG be.SBJV.3P.SG against nature-ACC.F.SG
(*De Officiis Ministrorum* 1.46.222, in Colish, 1990, p. 65)
'Decorum is life in accordance with Nature… and vice is what is contrary to Nature'

17. …*quoniam iuris natura fons sit* 'nature being the source of the law'.

Alongside constructions inherited from Hellenistic philosophy, new expressions typical of Christianity emerge, such as *secundum carnem* 'according to the flesh'. In Silver Latin *secundum* displays different collocational strength with *carnem* 'flesh' (log-likelihood = 98.565) versus *naturam* 'nature' (log-likelihood = 87.198), but this difference is not significant ($p = 0.404$, $df = 1$). In contrast, in Late Latin the difference in the collocational strength of *secundum* with *carnem* 'flesh' (log-likelihood = 711.828) and *naturam* 'nature' (log-likelihood = 71.192) is very highly significant ($p < 0.0001$, $df = 1$). This suggests that over time *secundum carnem*, indicating a truly Christian concept, has gained currency over *secundum naturam*, an expression of philosophical origin applied to a Christian concept.

The construction *secundum carnem* 'according to the flesh' has two meanings: one of conformity and one of limitation. In the first meaning, the expression contrasts with *secundum spiritum* 'according to the spirit'. The opposition between these two constructions represents a dualism which, as was explained above, Stoicism had completely eradicated and which specifically links early Christian thought to Plato's opposition between the soul and the body. In the *Bible*, this opposition is also found in the *Gospel according to Matthew* (26:41, "the spirit is willing but the flesh is weak").[18] This opposition is central to the thought of Paul the Apostle, who uses the Greek expression *katà sárka* [19] 'according to the flesh' in contrast with *katà pneûma* 'according to the spirit' (*Bible*, *Romans* 1:3–4 and *Galatians* 4:23, 4:29). The construction *secundum carnem* translates Greek *katà sárka*, much like *secundum naturam* translates *katà phýsin*. The expression *secundum carnem vivere*, clearly modelled on *secundum naturam vivere*, "denotes simply life lived at the level of decaying materiality" (Dunn, 1998, p. 65).

On top of the meaning of accordance discussed above, the construction *secundum carnem* has also a limitation meaning ('as far as the flesh is concerned, in flesh'), as in Example (17):

(17) *Quem…* ***secundum*** ***carnem*** *…* *hominem*
That.ACC.M.SG in.accordance.with flesh-ACC.F.SG man-ACC.M.SG
accipimus. (Novatian, *De Trinitate* 17, 3)
receive-PRS.1P.PL
'(Christ) whom we have received according to his human nature'

This construction contrasts with *secundum Deum* ('as far as God is concerned') and *secundum spiritum* ('as far as the spirit is concerned'), therefore forming

18. This concept in the Gospels links back to the notion of 'sinful flesh' in the Qumran/Dead Sea Scrolls (Ferguson, 2003, p. 529; Frey, 2006, p. 453).

19. The term sárx 'flesh' holds negative connotations in contrast with *sôma* 'body', which does not (see Dunn, 1998, p. 62–65 for a critical analysis of the term *sárx* in Paul).

a three-way opposition capturing the three hypostases of God in the doctrinal elaboration of the Trinity, commenced with the ante-Nicene fathers and fully formalised in the 4th century (Figure 8), as expressed for instance in the Nicene-Constantinopolitan Creed (381): *Credo in unum Deum, Patrem onnipotentem… et in unum Dominum Jesum Christum, Fiulium Dei unigenitum… et in spiritum sanctum, Dominum et vivificantem* 'I believe in one God, the Father almighty… And in one Lord, Jesus Christ, the only begotten Son of God… And in the Holy Spirit, the Lord and giver of Life.'

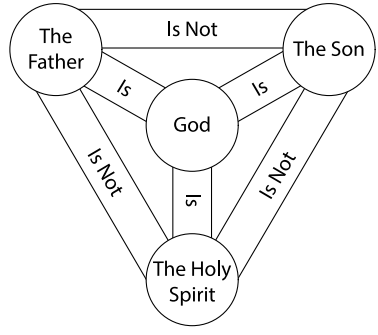


Figure 8. The Holy Trinity

The *secundum carnem* construction is one of the linguistic expressions[20] coined by Early Christian theologians to refer to one of the three persons of God. The concept of the Triune God is one of the theologically most complex notions in Christianity: there is no doubt that finding language to express this complexity was "a tenacious struggle" (Studer, 1994, p. 243). Language creativity and the adaptation/exploitation of older constructions as templates for new expressions were necessary resources allowing Christians to encode key concepts related to the Christian faith and identity. Among these concepts a prominent place was occupied by the mystery of incarnation, a notion which, as Ferguson (2003, p. 620) points out, was absolutely central to the Christian belief, especially in contrast with Judaism.

As a concluding remark, one must bear in mind the question of whether developments occurring in a specific genre – especially one quite removed from everyday speech such as philosophical/theological writing – could impact on 'everyday language' to the extent that they could be considered as propelling language change. One could in fact imagine that the contribution of constructions

20. Another expression is *secundum formam Dei/servi* 'in the form of God/the servant' (in my Early Medieval corpus there are four examples all by Isidore, Etymologies 7, 2, 46). I am not aware of any linguistic study centred on the expressions used to refer to the Trinity.

belonging to philosophy may be quite marginal overall in terms of constructional change.

This section aimed at showing the importance of the specific history of the *secundum* NP network of constructions. As Hilpert (2013b, p. 8) states, there are "pervasive cross-linguistic regularities […] but once the developments […] are studied in detail, it soon emerges that each construction has its own history."

9. Conclusion

This paper presented a quantitative analysis of the *secundum* NP ('according to' NP) construction as it developed in the history of Latin. This study focussed on the productivity and schematicity of *secundum* NP, which are shown to increase for centuries beyond the undocumented phase of grammatical constructionalisation/grammaticalisation.

Specifically, this paper has shown how *secundum* NP becomes more productive over time. The increase in its syntactic scope with conformity meaning between the Classical and Silver periods is compatible with the appearance of (inter)subjective meanings such as reportative evidentiality and attribution. The host-class expansion of the argument of *secundum* NP with a meaning of conformity also reveals a sudden increase during the Classical-Silver transition. This is explained with reference to Barðdal's (2008) model of productivity. Since constructions are the result of both cognitive and socio-cultural processes, the host-class expansion of *secundum* NP with conformity meaning is also understood as capturing and expressing salient cultural changes in the transition from Paganism to Christianity. Constructional networks for the *secundum* NP sub-construction are identified, proving that the conformity construction is actually less homogeneous and more dynamic than previously assumed. Attention is then paid to specific constructions (e.g. *secundum naturam* 'according to nature' and *secundum carnem* 'according to the flesh'), which prove to be idiosyncratic and culturally significant. My analysis shows that their usage patterns have an impact on the development of the *secundum* NP construction as a whole.

To conclude, my study supports the observation that constructional change includes frequency changes that are related to grammaticalisation without being exclusive to it. In fact, this study shows that some of these frequency changes extend well beyond the grammatical constructionalisation phase. These changes in the frequency of *secundum* NP are interpreted as (post-constructionalisation) constructional changes, on a par with semantic and syntactic changes. Importantly, these changes are understood in continuity with the process of grammatical constructionalisation. This points to the need for an increasingly tighter

integration between models of grammaticalisation and models of semantic and morphosyntactic change founded on diachronic constructional grammar, as exemplified for instance in Traugott & Trousdale (2013).

Abbreviations used in the interlinear glosses

1P First person
2P Second person
3P Third person
ABL Ablative
ACC Accusative
ACT Active
DEP Deponent
F Feminine
FUT Future
GEN Genitive
IMP Imperative
IND Indicative
INF Infinitive
M Masculine
N Neuter
NOM Nominative

PL Plural
POSS Possessive

PRS Present
PST Past
PTCP Participle

REL Relative pronoun
SBJV Subjunctive

SG Singular
. Functions of portmanteau morphemes or functions of morphemes whose segmentation is not shown
– Morpheme boundary
(…) Non overt element or sound change

References

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij, & J. van Marle (Eds.), *Yearbook of morphology 1992* (pp. 181–208). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-3710-4_7

Baayen, H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Baayen, R. H. (2009). Corpus Linguistics in Morphology. Morphological Productivity. In A. Lüdeling, & M. Kytö (Eds.), *Corpus linguistics. An international handbook*. Volume 2 (pp. 899–919). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110213881.2.899

Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.8

Barðdal, J. (2012). Predicting the Productivity of Argument Structure Constructions. *Berkeley Linguistics Society*, 32, 467–478.

Barðdal, J., & Gildea, S. (2015). Diachronic construction grammar. Epistemological context, basic assumptions and historical implications. In J. Barðdal, E. Smirnova, & L. Sommerer (Eds.), *Diachronic construction grammar* (pp. 1–51). Amsterdam: John Benjamins. https://doi.org/10.1075/cal.18.01bar

Butler, C. (1985). *Statistics in linguistics*. Oxford: Blackwell.

Bybee, J., & Thompson, S. (1997). Three frequency effects in syntax. Reproduced in J. Bybee (2007) (Ed.), *Frequency of use and the organization of language* (pp. 269–278). Oxford: Oxford University Press.

Bybee, J. (1985). *Morphology. A study of the relation between meaning and form*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.9

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612886

Bybee, J. (2003). Mechanisms of change in grammaticization. The role of frequency. In R. Janda, & B. Joseph (Eds.) *Handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell Publishers. https://doi.org/10.1002/9780470756393.ch19

Bybee, J. (2007a). Introduction. In J. Bybee (Ed.), *Frequency of use and the organization of language* (pp. 5–23). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.003.0001

Bybee, J. (2007b). *Frequency of use and the organization of language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Colish, M. L. (1990). *The Stoic Tradition from Antiquity to the Early Middle Ages. Volume 2: Stoicism in Christian thought through the sixth century*. Leiden: Brill.

Coussé, E. (forthcoming). Grammaticalization, host-class expansion and category change. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category change from a constructional perspective*. Amsterdam: John Benjamins.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803864

De Boer, E. (2005). *The gospel of Mary: Listening to the beloved disciple*. London: Continuum.

Ferguson, E. (2003). *Background of early Christianity*. Third Edition. Grand Rapids, Michigan: Eerdmans.

Ferrari, L. (1998). A gramaticalização de formas não-finitas como evidência da motivação conceptual do léxico. *Revista de Estudos Linguísticos Veredas*, 2(1), 103–115.

Gaeta, L., & Ricca, D. (2006). Productivity in Italian word formation: A variable-corpus approach. *Linguistics*, 44(1), 57–89. https://doi.org/10.1515/LING.2006.003

Goldberg, A. (1995). *Constructions. A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Guardamagna, C. (2016a). A cognitive-constructional approach to the development of Latin *secundum* NP. Unpublished PhD thesis. Lancaster University.

Guardamagna, C. (2016b). A diachronic semantic map for the Latin preposition *secundum* . *Journal of Latin Linguistics*, 15(2), 233–279.

Hammond, M. (1976). *Latin. A historical and linguistic handbook*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674332218

Heine, B. (2004). Grammaticalization. In B. Joseph, & R. Janda (Eds.), *The handbook of historical linguistics* (pp. 575–601). Oxford: Blackwell.

Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge: Cambrigde University Press. https://doi.org/10.1017/CBO9780511613463

Heine, B., & Narrog, H. (2011). Introduction. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 1–19). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199586783.001.0001

Hilpert, M. (2012). Diachronic collostructional analysis meets the Noun Phrase: Studying *many a Noun* in COHA. In T. Nevalainen, & E. C. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 233–244). New York: Oxford University Press.

Hilpert, M. (2013a). Corpus-Based approaches to Constructional Change. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 458–476). Oxford: Oxford University Press.

Hilpert, M. (2013b). *Constructional changes in English. Developments in allomorphy, word formation and syntax*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139004206

Himmelmann, N. (2004). Lexicalization and grammaticalization: Opposite or orthogonal? In W. Bisang, N. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization: A look from its fringes and its components* (pp. 21–42). Berlin: Mouton de Gruyter.

Hollmann, W. B. (2003). *Synchrony and diachrony of English periphrastic causatives: A cognitive perspective*. Unpublished PhD dissertation. University of Manchester.

Hopper, N., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambrigde University Press. https://doi.org/10.1017/CBO9781139165525

Kay, P., & Michaelis, L. (2012). Constructional Meaning and Compositionality. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 3 (pp. 2271–2296). Berlin: Mouton de Gruyter.

Kiss, T. (2007). Produktivität und Idiomatizität von Präposition-Substanti-Sequenzen. *Zeitschrift für Sprachwissenschaft*, 26(2), 317–345. https://doi.org/10.1515/ZFS.2007.022

Kühner R., & Holzweissig, F. (1966[1912]). *Ausführliche Grammatik der Lateinischen Sprache*. Second edition. Hannover: Hahnsche Buchhandlung.

Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume I. Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, R. W. (1991). *Foundations of cognitive grammar. Volume II. Descriptive application*. Stanford: Stanford University Press.

Lehmann, C. (1995 [1982]). *Thoughts on grammaticalization*. Munich: LINCOM (Revised second edition of *Thoughts on grammaticalization: A programmatic sketch*, 1982).

Lichtenberk, F. (1991). Semantic change and heterosemy in grammaticalization. *Language*, 67(3), 475–509. https://doi.org/10.1353/lan.1991.0009

Lindsey, W. M. (1915). *Short historical Latin grammar*. Second Edition. Oxford: Clarendon Press.

Poultney, J. W. (1980). The phonology of *-nd-* and the Latin gerundive. In H. J. Izzo (Ed.), *Italic and Romance. Linguistic Studies in Honor of Ernst Pulgram*. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.18.06pou

Rocha, L. F. (1998). Anàlise do caràter polissêmico do vocàbulo latino SECUNDUS. Universidade Federal de Juiz de Fora, Departimento de Letras, Lingüìstica – Grupo Gramàtica & Cognição. *Principia Caminhos da Iniciação Cientifica*, 3, 173–179.

Sharma, U., & Karma, S. K. (2003). *Western political thought. Volume 1: From Plato to Burke*. New Delhi: Atlantic.

Štichauer, P. (2009). Morphological productivity in diachrony: The case of the deverbal nouns in *-mento*, *-zione* and *-gione* in Old Italian from the 13th to the 16th Century. In F. Montermini, G. Boyé, & J. Tseng (Eds.), *Selected proceedings of the 6th Décembrettes* (pp. 138–147). Somerville, MA: Cascadilla Proceedings Project. http://www.lingref.com/cpp/decemb/6/paper2241.pdf

Studer, B. (1994). *Trinity and incarnation: The faith of the early church*. Edinburgh: T & T Clark.

Traugott, E. C. (1995). Subjectification in grammaticalization. In D. Stein, & S. Wright (Eds.), *Subjectivity and subjectivisation* (pp. 31–54). Cambridge: Cambrigde University Press. https://doi.org/10.1017/CBO9780511554469.003

Traugott, E. C., & Dasher, R. (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Vernon Arnold, E. (2014[1911]). *Roman Stoicism*. London: Routledge.

Vineis, E. (1998). Latin. In A. Giacalone Ramat, & P. Ramat (Eds.), *The Indo-European languages* (pp. 261–321). London: Routledge.

Zeldes, A. (2009). Quantifying constructional productivity with slot members. In *Proceedings of the NAACL HLT Workshop on Computational Approaches to Linguistic Creativity, June 5th, Boulder, CO* (pp. 47–54). Stroudsberg, PA: Association for Computational Linguistics. https://doi.org/10.3115/1642011.1642018

Zeldes, A. (2011). On the productivity and variability of the slots in German comparative correlative constructions. In M. Konopka, J. Kubczak, C. Mair, & U. H. Štìcha Waßnr (Eds.), *Grammar & Corpora, Third International Conference, Mannheim 22nd–24th of September 2009* (pp. 429–449). Tübingen: Narr.

Zeldes, A. (2012). *Productivity in argument selection. From morphology to syntax*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110303919

Latin corpus

The Latin Library corpus. Available at https://cqpweb.lancs.ac.uk/latinlib/ Accessed 28 February 2014.

Constructions at sentence level

The development of the conditional *caso* construction in Spanish

Anton Granvik
University of Gothenburg | HANKEN School of Economics, Helsinki

This paper analyzes the development of the abstract noun *caso* 'case' into a conditional marker in Spanish, as in *(en) (el) caso (de) que llueva, no saldremos* '(in) (the) case (of) that it rains, we'll stay in.' Building on diachronic corpus data I show that although the development of the conditional sense of *caso* in what I call the conditional *caso* construction can, in principle, be considered a textbook example of grammaticalization, in order to account for the formal variation which characterizes the construction as a whole, a construction grammar account might be preferable. Following Traugott & Trousdale (2013) I suggest that the development of the conditional *caso* construction can be seen as a case of constructionalization and subsequent constructional changes. From the perspective of Spanish *caso* the two approaches can be seen as complementary, rather than incommensurate, since they both highlight important aspects of the development of a new grammatical construction.

Keywords: Spanish *caso*, conditional conjunction, grammaticalization, constructionalization, constructional change

1. Introduction[1]

It is well known that languages develop grammatical elements such as (complex) prepositions and conjunctions out of (the combination of) lexical elements, such

1. This paper is part of the research project FFI2015-64080-P, ***Procesos de gramaticalización en la historia del español (V): gramaticalización, lexicalización y análisis del discurso desde una perspectiva histórica***, funded by the Spanish Ministry of Economy, Industry and Competitiveness. I would like to thank two anonymous reviewers as well as the editors of this volume for their valuable comments on a previous draft of this paper. The text has been significantly modified since then. All remaining errors and inconsistancies are my sole responsability.

https://doi.org/10.1075/cal.21.c8

as nouns, verbs and adverbs. In Spanish, many compound prepositions have the schematic format [prep + N + prep] or [adverb + prep], as *en frente de* [in front of] and *delante de* [before of] which can both be translated as 'in front of'.[2] Many complex conjunctions are very similar, with the addition of the subordinator *que* 'that': *en el caso de que* 'in the case that, if' and *siempre que* 'always that, if'. The transformation of the nouns forming the nucleus of compounds such as *en frente de* and *en caso de* into compound prepositions has been considered a typical case of grammaticalization (Meillet, 1975[1912]; Lehmann, 1985, 2002; Cifuentes Honrubia, 2003), since in their new function *frente* and *caso* have more grammatical meanings than as independent nouns. *Frente* can mean both 'forehead' and 'front', and both meanings are related to the idea of locative *beforeness*, whereas *caso* 'case', just as its English counterpart, has multiple senses such as 'event, happening; coincidence or chance; incident, occasion, situation', but also 'instance, case'.[3]

This paper focuses on the development of a new complex conjunction in Spanish based on *caso* 'case', namely *(en) (el) caso (de) que* 'in the case [of] that'. As far as I am aware, there exists no previous, detailed study of how the abstract noun *caso* developed into a conditional conjunction in Spanish, but Rosenkvist's (2004, pp. 162–171) analysis of the Swedish noun *fall* 'fall', which he compares to Latin *casus* and Greek *ptosis* 'a fall', provides a convincing explanation of the semantic change from a noun with the abstract meaning of 'case, instance' into a grammatical item indicating conditionality. In Swedish, *fall* is grammaticalized in combination with the preposition *i* 'in', i.e. *ifall* 'if', a development which is paralleled by Spanish *en caso*, Fr. *en cas*, It. *nel caso*, Pt. *em/no caso* and English *in case*.

Importantly for a grammaticalization account, Rosenkvist (2004, pp. 163, 169) considers that "the hypothetical meaning 'instance' is directly derived from the basic meaning 'fall'", and that "the possibility of a conditional interpretation is a latent property of *fall*." As we will see in the analysis of the data in Section 4, a similar development from an 'instance' meaning to a 'conditional' meaning can be identified in medieval Spanish.

An interesting fact about the Spanish complex conjunction involving *caso* is that it has not developed into a single form as is the case of Swedish *ifall* 'if'. Instead, the Spanish *caso* construction exhibits considerable formal variation both over time and at any given moment. On the one hand, there is the full form, *en el*

2. In fact, *en frente de* also has the meaning 'across from'.

3. The English terms in the text are translations based on the *Collins Spanish Dictionary* of the definitions in the Dictionary of the Royal Spanish Academy (DRAE, *s.v. caso*): "suceso, acontecimiento; casualidad o acaso; lance, ocasión o coyuntura". The last two English terms are taken from Rosenkvist (2004, p. 163).

caso de que 'in the case (of) that', including the initial preposition *en* 'in', the determiner *el*, the noun *caso*, and the preposition *de* 'of' before the subordinator *que*; on the other hand, there is the short form *caso que*, in which only the noun and the subordinator are present. As Examples (1a) and (1b) show, in present-day Spanish both forms have the same function of relating two propositions, one functioning as a condition for the realization of the other:

(1) a. pero eso no lo sabía el tigre y aun **en el caso de que** lo supiera, no daba signo alguno de detenerse ante tan inoportuno conocimiento.
(Víctor Cáceres Lara, *Tierra ardiente*, 1966)
'but the tiger did not know this and even in the case that it did know, it showed no sign of stopping in front of such an inconvenient piece of information.'

b. El sitio del combate solía ser algún paraje desierto, cercano a la línea divisoria entre dos condados, pues de este modo, **caso que** las autoridades del condado en donde tenía efecto la lucha intervinieran, se podía reanudar en el vecino. (Fabricio Valserra, *Pugilismo*, 1945)
'The combat used to take place in a deserted place close to the border between counties, because in this way, in case [that] the authorities of the county where the combat took place interceded, it could be retaken in the neighboring county.'

In between these two formally extreme variants, there are four more: *en el caso que, en caso de que, en caso que* and *caso de que*, since the elements *en*, *el*, and *de* can be either present or absent without any real change of meaning.

This paper has three main objectives. The first is to account for the development of the conditional meaning of the noun *caso* in what I call the Conditional *Caso* Construction (CCCxn): (*en) (el) caso (de) que*. In doing this I will compare how the well-established grammaticalization model of language change fares against the more recent constructionalization approach. Second, I will show that a constructionalist perspective can help explain the development of this construction by relating it to more abstract constructional schemas. This makes it possible to take into consideration the formal variation which characterizes the CCCxn. Finally, I will suggest that the CCCxn is partially schematic, and comes with several sub or allostructions.

In accordance with Construction Grammar (CxG), I assume that this complex expression, consisting of the grammaticalized noun *caso* and the subordinator *que*, as well as three 'optional' elements, constitutes a construction, in the sense that it can be conceived of as a conventional combination of form and meaning (cf. Croft, 2001; Goldberg, 1995, 2006; Traugott & Trousdale, 2013; Hoffmann & Trousdale, 2013). By adopting this theoretical perspective, the idea is that the

Construction Grammar view on language as a hierarchical network of more and less schematic vs. specific constructions provides an appropriate tool for addressing the formal changes which the construction undergoes after the establishment of the conditional meaning.

The paper is structured as follows. In Section 2 I introduce the corpus and methods used. Section 3 starts with a short discussion of the relationship between the grammaticalization and constructionalization frameworks. Section 3.2 then presents the compound prepositions and conjunctions in Spanish and relates their historical development to the grammaticalization and constructionalization models. Section 3 ends with a glimpse at Spanish conditional sentences (3.3). In Section 4 the main results of the corpus analyses are presented, and the development of the conditional *caso* construction is discussed from the perspectives of both grammaticalization and constructionalization. Some concluding remarks are presented in Section 5.

2. Corpus and methods

For the purpose of this study, I extracted 100 random cases per century of the sequences *caso que* and *caso de que* from the 279-million-word Corpus Diacrónico del Español (CORDE). With the earliest attestations found in the 13th century data, the initial sample included 716 cases of *caso que* (16 cases of *caso que* were found in the 13th century, with 100 cases more for the seven remaining centuries) and 383 cases of *caso de que* (in the 15th, 16th, and 17th centuries, 7, 36, and 40 cases were identified, to which 100 cases for the 18th, 19th, and 20th centuries are added). After manually analyzing the actual instances of the conditional conjunction *(en) (el) caso (de) que*, 536 cases could be identified. 272 of these cases correspond to the conditional conjunction *(en) (el) caso que* and 264 cases to *(en) (el) caso de que*. Table 1 shows the distribution of the six formal variants of the CCCxn included in the sample (i.e. the alternatives with the presence of *en, el* and *de*, respectively.

Although this is of course not the place for any actual analysis, it is interesting to observe that there are some differences when the variants with and without *de* before *que* are contrasted in Table 1. Before *caso que* the initial preposition *en* alone is used in 73% of the cases, whereas before *caso de que* both the simple preposition *en* and the preposition + determiner variant *en el* are almost as frequent (with 51% and 44%, respectively). Finally, Table 1 also indicates that when no initial preposition is used, then *de* is usually not used either (53 out of 67 cases of *caso que* represents 79%).

Table 1. The six different 'forms' of the Spanish conditional *caso* construction

Preceding elements	*caso que*	*caso de que*	Sum
en el	21 / 8%	134 / 51%	155
en	198 / 73%	116 / 44%	314
none	53 / 19%	14 / 5%	67
Sum	272 / 100%	264 / 100%	536

In Section 4, the corpus data is submitted to further diachronic analysis. First, the data is explored in order to see whether it is possible to identify how the conditional meaning came to be associated with *caso*. This development is addressed in terms of grammaticalization and constructionalization/constructional changes. Second, the 536 cases included in the sample presented above are used to illustrate how the different variants are distributed over time, with a focus on whether the presence of the elements *en, el* and *de* implicate a difference in meaning or usage of the construction. This analysis will concentrate on three main traits of the usage context of conditional *caso*: a) the conditional sentence type (if A, B; B, if A, or only if A), b) the verb form following the conjunction (past, present, or future subjunctive), and c) the kind of conditional reading (a possible, improbable, or impossible condition). The rationale behind these traits is presented in Section 3.3, and the results of the analysis in Section 4.2.

3. Theoretical background

3.1 Grammaticalization and constructionalization

Recent years have seen an increasing interest in addressing language change from a CxG perspective, resulting in the rise of so called Diachronic Construction Grammar (cf. Noël, 2007; Barðdal et al., 2015). One of the most influential recent works on the diachronic application of construction grammar is the monograph by Traugott & Trousdale (2013) (T&T from now on), which draws heavily on previous work on grammaticalization and lexicalization in an attempt to reformulate these models of linguistic change in terms of changes in constructions.

Traugott & Trousdale (2013) introduce two key concepts for the diachronic development of constructions, *constructionalization* and *constructional change*, which are defined as follows:

> Constructionalization is the creation of $form_{new}$–$meaning_{new}$ (combinations of) signs. It forms new type nodes, which have new syntax or morphology and new coded meaning, in the linguistic network of a population of speakers. It is accom-

> panied by changes in the degree of schematicity, productivity and compositionality. (T&T, 2013, p. 22)

> A constructional change is a change affecting one internal dimension of a construction. It does not involve the creation of a new node. (T&T, 2013, p. 26)

That is, whereas constructionalization affects both the meaning and the form parts of the sign, in constructional change only one of them is changed. As the first quote indicates, a construction is here understood as a form–meaning pairing which can be of any level of complexity (single or multiword unit) or schematicity (i.e. lexically specific or highly general).

In essence, what T&T do is to use the three notions of (changes in) schematicity, productivity and compositionality to create a model which accommodates changes previously considered cases of both grammaticalization and lexicalization (cf. Brinton & Traugott, 2005). As such, even cases of degrammaticalization, which have posited problems for the claimed unidirectionality of grammaticalization, can be included in the constructionalization model (cf. Trousdale & Norde, 2013).

Since the focus of this paper, the conditional *(en) (el) caso (de) que* construction, involves the creation of a grammatical function/meaning, describing this development as a case of grammaticalization seems intuitively very natural and feasible. This is why, in accordance with Heine et al. (2016), who argue that "in spite of a number of similarities … works on constructional change and grammaticalization are not based on the same theoretical perspective and, hence, differ from one another in the way the development from compounding to derivation is analyzed", I will first try to account for the development of conditional *caso* in terms of grammaticalization.

As we will see, the four grammaticalization parameters presented in Heine et al. (2016, p. 151), context extension, desemanticization, decategorialization, and erosion, can be used to account for the development of the conditional meaning of *caso* quite well. My analysis will also involve the notions of bridging and switch contexts introduced by Heine (2002). Following Heine (2002, p. 86), in a bridging context "[t]here is a specific context giving rise to an inference in favor of a new meaning," whereas in switch contexts "[t]here is a new context which is incompatible with the source meaning." The final stage in Heine's model is conventionalization, which applies to a situation where "[t]he target meaning no longer needs to be supported by the context that gave rise to it: it may be used in new contexts" (2002, p. 86). However, it is important to keep in mind that the different stages in Heine's (2002) model need not be exclusive, but may coexist at any point in time.

Following their distinction of grammaticalization from constructional changes, Heine et al. (2016) make the case that their model of grammaticalization

is more apt for describing the development of a series of related grammatical meanings associated with a (former) morphologically nominal element *ví* 'child' in Ewe. In this paper, however, I will make the opposite choice, and suggest that the formal variation within the development of *(en) (el) caso (de) que* can best be apprehended by adopting a CxG approach. This idea will be developed a step further in the next subsection, where the paradigm of Spanish compound prepositions and conjunctions is introduced.

3.2 The Spanish compound prepositions and conjunctions

The creation of compound prepositions and conjunctions in the Romance languages as compared to their common ancestor Latin has been considered a paradigm case of grammaticalization ever since Meillet's (1975[1916]) famed article on the renewal of Latin conjunctions (cf. Lehmann, 2002; Cifuentes Honrubia, 2003; Fagard & De Mulder, 2007; Granvik, 2013, 2014; Sansò & Mauri, 2014). An interesting fact about the renewal of conjunctions is that most new forms are complex, originally multi-word units. In Cano's (2010, pp. 150–151) words: "in the whole history of Spanish not a single one-word unit has been incorporated into the inventory of subordinating connectives. On the contrary many have disappeared."[4]

It is not surprising, then, that Lehmann (2002), for example, uses the development of both compound prepositions and subordinating conjunctions in Spanish as an example of the intimate relationship between grammaticalization and lexicalization. Expressions such as *hacia* 'towards' and *a pesar de que* 'in spite of' constitute clear cases of lexical items, such as Latin *face* 'face' and Spanish *pesar* 'sorrow, regret' developing into grammatical items.

Lehmann (2002, p. 11) makes the point that the transformation of *face ad* into the preposition *hacia* 'toward' includes reanalysis of the complex pattern so that it is considered a single unit, i.e. lexicalization. Similarly, a complex conjunction such as *a pesar de que*, which ends up "containing" the subordinator and thus "takes a clause as its complement" (Lehmann, 2002, p. 13), also ends up as part of the inventory of Spanish subordinating conjunctions. In Lehmann's (2002) account of the creation of complex grammatical expressions such as *hacia* and *a pesar de que*, the grammaticalization and lexicalization processes are mutually complementary, and support one another. This idea is taken up by Girón (2008, p. 34), who argues that "lexicalization and grammaticalization appear and develop

4. My translation. The original reads: "en toda la historia del español no se ha incorporado una sola unidad simple al elenco de nexos de subordinación. Por el contrario, varias han desaparecido."

in tight connection: a grammaticalization always requires a previous lexicalization, and has as a result another lexicalization which is more stable than the first."[5]

Using different terminology in his grammaticalization analysis of the German preposition *Richtung* '(in the) direction (of)', Rostila (2004, 2006) argues that the storage of the grammaticalized expression is essential for subsequent changes. A very similar idea is also found in T&T (2013), who conceive of (grammatical and lexical) constructionalization as being both preceded and followed by constructional changes of different kinds: "When grammaticalization occurs is key to identifying when new grammatical form$_{new}$–meaning$_{new}$ pairing occurs. What precedes are enabling constructional changes, what follow are constructional changes" (T&T, 2013, p. 148). Here their use of the term "pairing" seems to imply storage.

Apart from the above examples, the history of Spanish has seen the creation of hundreds of complex prepositions, and (subordinating) conjunctions. The sources for these are highly variable, but include nouns, adjectives, verbs and participles, adverbs, etc. (see Cifuentes Honrubia, 2003, pp. 94–103). Although the grammatical status of these word classes is much disputed within Spanish grammar,[6] the relationship between adverbs, prepositions and conjunctions is, in fact, relatively systematic. In what follows, I will attempt to characterize this systematic relationship with the aid of constructional schemas.

In present-day Spanish, adverbs are generally the least complex of the three kinds of units and are mostly conceived of as single words, compound prepositions occupy a middle ground and conjunctions are typically complex (even com-

5. The original Spanish passage reads: "lexicalización y gramaticalización surgen y se desarrollan en estrecha cooperación: una gramaticalización requiere siempre una lexicalización previa y tiene como resultado una segunda lexicalización más estable que la primera" (Girón Alconchel, 2008, p. 34).

6. Many authors prefer the label *partículas* 'particles' as a cover term for adverbs, prepositions and conjunctions, under the license that they are invariable elements which establish relationships between other sentential elements and which are typically related with adverbials (Pavón Lucero, 1999, p. 567). The main reason for the problematic status of the particles is that there is no agreement on how to meaningfully distinguish a) coordinating and subordinating conjunctions from one another; b) subordinating conjunctions from prepositions and adverbs; and c) prepositions and adverbs from one another (see Devís Márquez, 1994; Pavón Lucero, 1999; Garrido Vílchez, 2008; NGLE § 29.9, § 31.11, Granvik, 2013).

positional). [7] Accordingly, there are three main schemas that can be used to create grammatical compounds in Spanish:

i. an adverb(ial) can be made into a preposition by adding a prepositionalizer (most often the preposition *de*) to it (see Bartens & Granvik, 2012).[8]
ii. many adverbs can be turned into (subordinating) conjunctions by adding the subordinator *que*.
iii. many prepositions can also be made into (subordinating) conjunctions by adding the subordinator *que* (see Granvik, 2013).

These schemas, which are exemplified in Table 2, have been active during the history of Spanish and, to a significant extent, still are.

Table 2. Constructional schemas for (creating) compound prepositions and conjunctions in Spanish

Constructional schema		Examples
i.	adverb + *de* > compound preposition[9]	*arriba, antes, debajo, cerca, aparte… > arriba de, antes de, debajo de, cerca de, aparte de…* 'above', 'before', 'under', 'near, close', 'separately' > 'on top of', 'before', under', 'close to', 'apart from'
ii.	adverb + *que* > subordinating conjunction	*aún, siempre, maguer > aunque, siempre que, maguer que…* 'still', 'always', 'I'd wish' > 'despite, although', 'if, as long as', 'despite, although'
iii.	preposition + *que* > subordinating conjunction:	*por, con, antes de, aparte de, a fin de, en caso de… > porque, conque, antes de que, aparte de que, a fin de que, en caso de que…* 'for', 'with', 'before', 'apart from', 'in order to', 'in case of' > 'because', 'since, if', 'before', apart from [the fact] that', 'in order to + that', 'in case [of] that'

7. Indeed, many grammarians do not consider the conjunctions formed by adding *que* to a (complex) preposition 'syntactic units'. The *Nueva Gramática de la Lengua Española* (NGLE, § 31.11l), for example, prefers the (syntactically motivated) view that the complex conjunctions such as *en caso de que* do not constitute conjunctions per se, but rather are complex prepositions that govern sentential complements introduced by the complementizer *que*.

8. Lehmann (2016) calls this function *relationalizer*.

9. *A* is also frequently used as a final element in Spanish compound prepositions, e.g. *junto a, rumbo a*, but *de* is undoubtedly the most typical one. It also needs to be noted that there are many more patterns than [adverb + *de*] for creating compound prepositions in Spanish, e.g. prep+N+prep, adj+prep, etc. Cifuentes Honrubia (2003, pp. 114–115) lists a total of 17 patterns. The prep+N+prep pattern is of particular importance for this paper since it yields the compound preposition *en caso de*.

Although this is a highly simplified account, it captures several relevant characteristics of this paradigm in Spanish. In the first place, it emphasizes the fact that the schemas [adv + *de* > prep], [adv + *que* > conjunction] and [prep + *que* > conjunction] are and have been productive in Spanish. Second, it underlines the importance of the elements *de* and *que* for changing the category of the compound (see Bartens & Granvik, 2012; Granvik, 2013; Lehmann, 2002, 2016). Third, it illustrates the fact that the compound prepositions and conjunctions are not closed inventories, but rather open classes that can be continuously added to (see Cifuentes Honrubia, 2003; Hoffmann, 2004; Codita, 2013). Fourth, from the perspective of Construction Grammar, and considering that the subordinating conjunctions are formed systematically following the schemas ii) and iii) above, an instance of this schema such as *(en) (el) caso (de) que* can be considered a micro-construction (cf. T&T, 2013, p. 14).

3.3 Spanish conditional sentences

The conditional sentences have undergone significant changes from Latin to the Romance languages, including a restructuring of both the verb forms used in the protasis and apodosis parts, and the explicit marking of different kinds of conditional sentences (see Marcos Marín, 1979; Penny, 2002; Garrido Sepúlveda, 2015). Of interest for this paper is the fact that, where most present-day accounts of conditional sentences include three types of conditionals, i.e. possible (real), improbable (potential) and impossible (irrealis), which can all refer either to the present/future (i.e. non-past] or the past, in medieval/classical Spanish there were no formal means of distinguishing between improbable and impossible conditions, past or present (Garrido, 2015, p. 195). Apart from the possible (real) conditions, which are characterized by indicative verb forms in both Latin and Spanish, the exact reading of improbable (potential) and impossible (irrealis) conditional sentences must be made taking into account contextual factors. This situation is depictured in Table 3.

Table 3. Structure and form of improbable and impossible conditional sentences in medieval Spanish (apud Penny, 2002, p. 251)

Medieval Spanish	Past	Non-past
Improbable	1. *Si pudies(se), fizieralo*. 'If he was able to, (but I don't think he was), he would have done it'	2. *Si pudies(se), fizieralo.* 'If he were to be able, (but I don't think he is/will be), he will/would do it'
Impossible	3. *Si pudies(se), fizieralo.* 'If he had been able to, (but he wasn't), he would have done it'	4. *Si pudies(se), fizieralo.* 'If he were to be able, (but he definitely isn't/won't be), he would do it'

This system was slowly modified, and by the 17th century the compound preterit forms (formed with auxiliary *aver* 'to have') had entered the paradigm, enabling a distinction to be made between past improbable and impossible conditions: *si pudiese* 'if he could' and *si hubiese podido* 'if he had been able'. This contrast between simple and compound past subjunctives (the *hubiese podido* form would enter cell 3 in Table 3), however, does not enable distinguishing between past and non-past improbable conditions (cells 1 and 2), and non-past improbable and impossible conditions (cells 2 and 4).

Despite the significant changes between Latin and Spanish, one essential thing remains the same: the default conditional marker is *si* 'if' even in present-day Spanish. However, over time several alternative conditional markers would evolve, such as *a/con condición (de) que*, *(en) (el) caso que*, *siempre que, con tal (de) que*. Mondéjar (1966) lists 14 groups of complex conditional conjunctions, one of which is based on *caso.* According to Mondéjar (1966, pp. 235–236), there are differences in the nuances that the conditional marker can introduce in the presentation of the condition. So, with *si* he considers that the condition is presented in a neutral fashion, whereas with *en caso de que* in a sentence such as *En caso de que vinieras a casa a recogerme, te acompañaría* 'In case you came to get me at home, I would accompany you', reference is made to "a fact whose realization is relatively improbable" (1966, p. 236). Julián (2009, p. 618), on the contrary, considers that with *en caso que* "it is understood that the fulfilling of the condition is eventual, possible."

Given these contradictory evaluations of the semantic import of *en caso de que* on the conditional sentence, on the one hand, and the fact that, contrary to *si* 'if', the compound conditional conjunctions are mainly followed by a subjunctive verb form, on the other, the kind of conditionality expressed by this complex conjunction cannot be established a priori. However, given the semantics of the noun *caso* 'instance, occasion', it seems plausible that, at least originally, this form was used to express possible (real) conditions (see Examples (2) and (3) below). If this can be observed to be the case, then a change in the kind of conditionality relation that the CCCxn expresses (e.g. from possible to either improbable or impossible conditions) might be seen as evidence of a further semantic development.

These aspects regarding the kind and types of conditionality and conditional sentences will be investigated in the corpus analysis in Section 4. Specifically, I will analyze the contextual uses of conditional *caso* from the point of view of the following three questions:

a. To what degree do they exemplify the default pattern [if A, B], an inverted one [B, if A], or a single-member [if A] condition?
b. What verb forms (subjunctive and indicative) are found in the protasis?

c. Which different kinds of conditional readings are expressed: real, possible, or impossible?

As indicated above (Section 2), these criteria will be used as descriptive means for comparing and contrasting the different stages and formal variants of the construction. Although, the first two criteria are reasonably easy to determine objectively by observing the actual use of conditional *caso*, with regards to (c) a careful reading of each corpus example is needed.

4. The conditional *caso* construction in use

Despite the undeniable status of *(en) (el) caso (de) que* as a conditional conjunction in Spanish, it remains to be seen to what extent the evolution of the conditional meaning can be observed in the corpus data. In Section 4.1 I will explore the available corpus data in order to account for the development of the conditional meaning of *caso*. I will then shortly evaluate how this development can be accounted for by the grammaticalization and constructionalization models. In Section 4.2 the emphasis will lie on discussing the implications of the formal variation within the CCCxn for its constructional status. Section 4.3 offers a discussion of the appropriateness of the grammaticalization and constructionalization models for accounting for the development of the CCCxn as a whole.

4.1 The development of a conditional

In this subsection the focus lies on the 536 conjunctional uses of *caso* retrieved from CORDE. The diachronic distribution of the six variants is presented in Table 4, which reveals considerable diachronic variation in the form(s) of the conjunction. Specifically, the figures in Table 4 show that almost two thirds of the cases of conditional *caso que* without *de* are found between 1300 and 1599 (176/272 = 65%), while as much as 83% (218/264) of the cases of conditional *caso de que* are found between 1700 and 1974. These findings suggest that it is in the combination *en caso que* where the conditional reading originates. They also indicate that the history of the CCCxn does not only involve the development of a new conditional meaning in the form *en caso que*, but also a change in the form of the conjunction. Before the 1600s, the predominant form is *en caso que*, whereas the forms with *de* predominate from 1700 on. Although the determiner *el* is first attested in the 14th century, its usage increases significantly in the 18th century, and in the last two centuries *en el caso de que* is the most frequently used variant.

Table 4. Distribution of the different variants of the conditional *caso* construction from the 1300s to the 1900s

	1300s	1400s	1500s	1600s	1700s	1800s	1900s	Total
caso que	6	22	7	9	7	1	1	53
en caso que	68	30	33	26	28	5	8	198
en el caso que	1	8	1	1	4	5	1	21
Partial total	75	60	41	36	39	11	10	272
caso de que	0	0	1	3	1	7	2	14
en caso de que	0	2	17	20	41	20	16	116
en el caso de que	0	1	2	0	29	50	52	134
Partial total	0	3	20	23	71	77	70	264
Total	75	63	61	59	110	88	80	536

As shown in Table 4, the use of *caso* as center of the conditional conjunction is first attested in Spanish in the 14th century. As Examples (2) and (3) reveal, the conditional meaning is clear both with and without initial *en*. In both examples the past subjunctive form is used in the protasis, *qujsiesen* '[if] they wanted' and *viniesse* '[if] he came', suggesting a hypothetical reading of the condition. However, in neither case does the condition seem to be impossible, despite the past subjunctive forms.

(2) Ca **en caso que** del qujsiesen lleuar el auer que esto era cosa que non podia ser
(*Crónica de 1344*, 1344)[10]
'Since in case they wanted to take his possessions, then this was something that could not be.'

(3) Et d'alli auant los fidalgos, quando fazien homenage alguno de alguna fortaleza que les dassen encomanda, lo fazien con condicion que **caso que** el rey hi viniesse lo deuiessen tan tost recullir.
(Juan Fernández de Heredia, *Gran crónica de España*, III, 1376–a 1391)
'And from then on the noblemen when they paid homage to any given fortress that gave them a trust, they did it with the condition that in case [that] the king came there, they should collect as soon [as possible].'

Although Examples (2) and (3) are illustrative of the new conditional conjunction, they tell us little about how the conditional meaning came to be associated with the noun *caso*. As stated above, the origins can be traced to the meaning of *caso*, and its use to refer to situations of different kinds (cf. Rosenkvist, 2004).[11] More

10. When the author is stated as anonymous in CORDE, only the title is mentioned.

11. Schmid (2000, pp. 289–291) uses the term 'contingency' to characterize conditional shell noun uses in English, of which *case* is one.

importantly, when combined with locative *en*, in e.g. *en tal caso* 'in such [a] case', the stage is set for making inferences which approach the conditional. It is quite easy to find examples of this kind in texts from the 13th century:

(4) Ca **en tal caso** como este ssy ffirmasse el vn testigo que ffuera en vna ora. & el otro en otra non deue valer ssu testimonjo. (*Espéculo de Alfonso X*, a 1260)
'Since in a case such as this if one witness would sign that he was there at one time, and another at another his testimony must not be valid.'

(5) E porende el fijo non deue casar con la madre njn la fija conel padre njn avn deue casar conlos otros parientes muy çercanos Sacando con dispensaçion & en algund caso donde se podiese seguir muy grand bien E **ental caso** podiese otorgar lo que es comonal mente negado alos otros. (*Castigos*, 1293)
'For that reason a son must not marry the mother nor the daughter with the father and neither must they marry with other close relatives except with dispensation and in some case where there might follow good things, and in such [a] case one could grant what is commonly denied to others.'

In (4) *en tal caso* is followed by conditional *si* 'if' (here in its medieval spelling *ssy*), so it is easy to see that *en tal caso* identifies the situation which acts as the starting ground for the actual condition expressed by *si*: that is, in a situation such as this one, if 's/he did something then something else would also apply'. In (5) the conditional reading of *caso* is more immediate, since here *en tal caso* is directly followed by the past subjunctive *podiese* 's/he could'. *Caso* thus directly sets the stage for the possible/hypothetical meaning of the subjunctive: 'in such a case one could do something'.[12]

Three further examples from the 14th century indicate how the noun *caso* can carry the meaning 'possible situation' also in the form *en caso que*. Representing a further step compared to the sentences in (4) and (5), Examples (6) to (8) can be considered bridging contexts (Heine, 2002) for the grammaticalization of *caso* as a conditional marker in the conjunction format, since in these sentences the conditional meaning is clearly available, despite not being entirely certain.[13]

12. Marcos Marín (1979, p. 102) calls *en tal caso* a "conditional formula" ("fórmula condicional"). According to Schmid (2000, p. 291) *in this case* in English can also be used to refer "back to the information given in [a] preceding sentence and [shell] it as PROTASIS."

13. Note that all three examples are found in the same late 14th century document, *Tratado de la comunidad*, and they are not, really, chronologically earlier than the first conditional uses of *en caso que*. However, the fact that they exist and are of such an early date as the 1370s is in and by itself significant, since diachronic data are seldom as precise as researchers would desire.

(6) **En caso que** la justiçia prinçipalmente sea neçesaria al regimiento de la comunidat, así mesmo es muy conplidero al bien de aquella concordia e unidat. (*Tratado de la comunidad*, c 1370)
'In case [that]/in the case where justice is mainly necessary for the government of a community, it is also very positive for the well-being of that concord and unity.'

(7) E por el contrario, quando el estómago del cuerpo humanal, **en caso que** tenía mucha vianda, si non la destribuye segunt que deve, e la enbía e reparte a los mienbros segunt que le es menester, los dichos mienbros se enflaquecen e adelgazan. (*Tratado de la comunidad*, c 1370)
'On the contrary, when the human stomach, in case it has much meat, if it does not distribute it as it must and send out and distribute it to the body parts as it needs to, said body parts become weak and skinny.'

(8) E los buenos casados deven de bevir sienpre en caridat **en caso que** sean dos cuerpos, sea una voluntad e bivan bien a Dios e al mundo. (*Tratado de la comunidad*, c 1370)
'And the well wed must live always in charity [;] in case they/there are two bodies, let there be one wish; and may they live well in front of God and the world.'

In (6), there is the 'case that' *justice* may be necessary for the community, but this part of the sentence could also be paraphrased as 'In the situation in which justice is necessary', which, of course, implies a relative clause reading of *(caso) que*. Here the sentence initial position and the present subjunctive form *sea* make it more probable that we are dealing with a complement clause with conditional nuances. In (7) *en caso que* is used after a temporal clause introduced by *quando* 'when', and it is combined with an indicative, so the relative clause, *caso* as 'situation' reading prevails: 'in a case where it had much meat in it'. This reading is further enforced by the fact that *en caso que*-clause is followed by a conditional sentence with *si*. Example (8), finally, is another borderline case, where the double subjunctive forms, *en caso que sean dos cuerpos* 'in case there are two bodies' and *sea una voluntad* 'let there be one wish', make a conditional reading more probable, although the situational reading remains present. Compared to (6) to (8), the sentences in (2) and (3) (above) can be considered switch contexts (Heine, 2002) for the development of the conditional meaning, since they are undoubtedly conditional (they conform to the typical [if A, B] conditional sentence pattern and include past subjunctive forms). In (2) and (3), then, the most natural paraphrase would be 'if they wanted to take' and 'if the king came there', respectively.

Although these examples clearly show that *caso* develops a conditional meaning, and that *en caso que* grammaticalizes into a conditional conjunction, there are a couple important points to be made in order to complete the picture.

First, the conditional meaning seems to arise in combination with *en* 'in', with the adverbial *en (tal) caso* setting the stage for the conditional reading (cf. Schmid, 2000, p. 290; Rosenkvist, 2004, p. 165). With this pattern established,[14] adding the subordinator *que* in order to establish the conditional conjunction is a natural step. When *que* is added following the [adv_*que*] schema, there is less need for the determiner *tal*, so its absence from the conjunction *en caso que* (e.g. Example (2)) is not surprising. The importance of initial *en* is also enhanced by the fact that examples indicative of bridging contexts for the conditional meaning are only found with *en caso que* (cf. Examples (6) to (8)), whereas *caso que* (as in (2)) seems to already represent a stage of conventionalization of the conditional meaning.

Second, the early appearance of the form *caso que* also merits a comment. As the figures in Table 4, above, indicate, *caso que* reaches significant usage frequencies a little later than *en caso que* (6 vs. 68 cases in the (second half) of the 14th century, 26 vs. 30 in the 15th), so it seems reasonable to suppose that it was created on the basis of the first by omitting the initial preposition.[15]

An explanation for why initial *en* was omitted can be found in Rostila's (2004, 2006) account of the grammaticalization of the German preposition *Richtung*. Following Rostila (2006, p. 10), once *en caso que* is established as "a meaning unit that can be manipulated as a whole, […] the atomic and specific construction/lexical item *in* is not needed in its processing and can be reduced." This way, the short form *caso que* has the typical two-word conjunctional format (*cf. aunque, maguer que*, etc. see Bartol, 1986; Ibba, 2007). This order of events also explains why *caso que* is not found in ambiguous conditional contexts, whereas *en caso que* include several such examples. Instead, the reduced form *caso que* seems to function as conditional conjunction from the beginning.

In view of the two models of grammatical change that constitute the theoretical basis for this paper, the changes observed so far can be resumed as follows. From a grammaticalization perspective, conditional *en caso que* clearly involves desemanticization: *caso* changes from an abstract noun with the meaning 'instance, situation' to a grammatical element expressing 'condition'. *(En) caso que* also undergoes context extension, since the use of conjunctional *(en) caso que* is clearly different as compared to nominal *caso* in e.g. the adverbial expressions *en*

14. Out of the 30 cases of *caso* found in the 13th century, 23 are instances of the pattern *en* + determiner + *caso*.

15. Sansò & Mauri (2014, p. 113) document the appearance of *caso che* in Italian in the 16th century, and consider it to derive from the prepositional version *in/nel caso che*.

tal caso, en este caso. The grammatical meaning and functioning as a conjunction are also indicative of a decategorialization of the noun. The only grammaticalization parameter that does not seem to be involved is erosion, but it needs to be noted that Heine et al. (2016, p. 151) consider erosion a tendency and not an actual parameter or principle of grammaticalization.

From a constructionalization perspective, T&T (2013, p. 22) note that the creation of a new construction involves changes in schematicity, productivity and compositionality. As the figures in Table 5 indicate, the new *en caso que* construction is increasingly productive in its new function as a conditional conjunction.[16] In the sense that the conjunctional function is more schematic (i.e. abstract) than the referential use of *caso* as a noun, also this criterion is met. In addition, the coexistence of *en caso que* and *caso que* can be interpreted as the creation of a schematic construction with two members: a more chunk-like, grammaticalized one, *caso que*, and another, more compositionally accessible one, *en caso que* (cf. T&T, 2013, pp. 116–117). From this perspective, one can see how the development of more chunk-like *caso que* out of *en caso que* involves a more abstract level of representation influencing the particular instance, showcasing a loss of compositionality.

Table 5. Relative frequencies (per million words) of the two main variants of the CCCxn from 1300 to 1974

	1300s	**1400s**	**1500s**	**1600s**	**1700s**	**1800s**	**1900s**
caso que/million	0,059	0,085	0,075	0,054	0,074	0,003	0,002
caso de que/million	0	0,001	0,003	0,005	0,123	0,136	0,088

Although the development of the conditional meaning of *caso* can be characterized both as a case of grammaticalization and as a (grammatical) constructionalization, in view of the data presented so far, the three relevant grammaticalization parameters (context extension, desemanticization and decategorialization) appear more directly observable and relevant when it comes to the development of the conditional meaning. This is perhaps so, because the development of a grammatical meaning from a referential lexical element (*caso*) within a certain grammatical pattern is a paradigm case of grammaticalization. In this first step of the development of conditional *caso*, the question of schematicity, productivity and compositionality are less immediately important. However, as was illustrated above regarding the relationship between *en caso que* and *caso que*, when

16. From the 17th century on, also the *en caso de que* variant becomes increasingly frequent, but this aspect does not concern the present discussion.

the subsequent steps of the development of the CCCxn are considered, i.e. the creation and coexistence of several formats of the CCCxn, the constructional factors become more important. This will become evident in the following subsections.

4.2 Formal variation and constructional status

In this subsection I will address the further development of the CCCxn, that is, the addition of more elements to the construction. A central aim is to elucidate to what degree the different forms translate into differences in meaning or usage. Following Goldberg (1995, pp. 67–69), and cognitive linguistics, in general, a change in form is expected to reflect a difference in meaning (or pragmatics). In doing this I will advance chronologically, and go from the minimal towards the maximal, contrasting first the medieval forms *caso que* and *en caso que*, then the importance of the introduction of *de* before *que* in the forms *en caso que* and *en caso de que*, and finally the main alternation between *en caso de que* and *en el caso de que* in the so-called modern period, from the 18th century on.

As stated in Section 3.3, I will address the question of differences in meaning and usage in terms of three factors: conditional sentence type, verb form in the protasis, and kind of conditional reading. The conditional sentences are divided into three formal types: typical [if A, B], inversed [B, if A] and one-member-conditionals [if A]. The verb forms that can be used in the protasis are, obviously, quite numerous, so for ease of exposition and also in order to assure (a certain degree of) quantitative reliability, I limit the analysis to three subjunctive forms (past, present and future) as well as the indicative.[17] The kind of conditional meaning, finally, is based on the traditional three-way distinction between possible (real), improbable (potential) and impossible (irrealis) conditions (see Penny, 2002 and Garrido Sepúlveda, 2015). The kind of conditional meaning was determined separately for each case, based on a detailed reading of the context of usage, but it must be acknowledged that this classification involves a certain degree of subjective judgement.

4.2.1 *The early variation: En caso que vs. caso que from 1300 to 1499*

The contrast between *en caso que* and *caso que* in the late medieval period is presented in three tables, Tables 6 to 8. Judging by the X^2 tests I have performed, overall there seem to be few semantic and functional differences between the two early alternatives. The only significant result can be found in Table 8, which reveals that

17. The quantitative motivation lies in the fact that using and including all attested verb forms in the tables would lead to a high number of cells with very low counts. This, of course, hinders statistically reliable testing even at the basic level of X-square tests.

caso que is used to a certain extent to express improbable conditions, and even one case of an impossible condition, whereas *en caso que* is limited to possible conditions. This preference is probably at least partially reflected also in Table 7, where *caso que* shows a higher frequency of use with the past subjunctive (50% vs. 36% for *en caso que*). This is the only form which could be used to express improbable and impossible conditions in the medieval language, although it was used for possible conditions as well.

Table 6. Distribution of the early variants *en caso que* and *caso que* in different conditional sentences types (1300–1499)

	en caso que	***caso que***	**Sum**
if A, B	66 / 67%	20 / 71%	86 / 68%
B, if A	20 / 20%	8 / 29%	28 / 22%
if A	12 / 13%	0 / 0%	12 / 10%
Sum	98 / 100%	28 / 100%	126 / 100%

X-squared = 4.1348, *df* = 2, *p-value* = 0.1265.

Table 7. Distribution of different verb forms in the protasis after *en caso que* and *caso que* (1300–1499)[18]

	en caso que	***caso que***	**Sum**
past subj	35 / 36%	14 / 50%	49 / 39%
pres subj	40 / 41%	7 / 25%	47 / 37%
fut subj	8 / 8%	1 / 4%	9 / 7%
indicative	15 / 15%	6 / 21%	21 / 17%
Sum	98 / 100%	28 / 100%	126 / 100%

X-squared = 3.736, *df* = 3, *p-value* = 0.2914.

This finding can be interpreted as indicating that the fixed conjunction (*caso que*) is more grammaticalized, in the sense that it enters the domain of improbable conditionality, whereas *en caso que*, being a little more semantically and structurally transparent does not. Another indication of a more specialized conditional function for *caso que* is illustrated in Example (9), where it is followed by a plu-

18. The asterisk in front of the X^2 values calculated for some of the tables indicates that the test is not reliable, i.e. that the "Chi-squared approximation may be incorrect". This is due to the fact that several tables include cells with very low counts, and the Chi-square test is sensible to this. However, given that most of the p-values are above the significance level anyway, this unreliability is perhaps less of a problem than it might have been had the p-values been significant.

Table 8. Different kinds of conditional readings for the early variants *en caso que* and *caso que* (1300–1499)

	en caso que	*caso que*	**Sum**
possible	78 / 100%	16 / 57%	94 / 89%
improbable	0 / 0%	11 / 39%	11 / 10%
impossible	0 / 0%	1 / 4%	1 / 1%
Sum	78 / 100%	28 / 100%	106 / 100%

* *X-squared* = 37.696, *df* = 2, *p-value* = 6.522e-09.

perfect indicative. Cases with a pseudo-conditional reading such as this one are not attested in my sample with *en caso que*. On the contrary, in this period *en caso que* is found in examples with no apodosis (see Table 6) and where the conditional meaning is borderline at most (cf. Examples (6) to (8), above).

(9) le mandaua que viniese a el ally a Madrid e **caso que** le auia herrado en algunas cosas, que el no paraua mientes a ello (*Gran crónica de Alfonso XI*, 1348–1379)
'he [King Alfonso] ordered him [the master] to come to him and [in] case that he had failed him in some things, he would pay no mind to that.'

4.2.2 *Introduction of the [prep_que] variable: En caso que vs. en caso de que from 1500 to 1799*

The preposition *de* is first attested in the CCCxn in the 15th century, but it only becomes frequent in the 16th century. The introduction of the preposition *de* must be related to the spread of a new syntactic form which starts in the 16th century, namely what Pountain (2014, p. 52) has called the first wave of the variable [prep_*que*]. As is well known from present-day Spanish, the [prep_*que*] variable is found with all kinds of clausal complements, including those with verbal, adjectival, nominal, adverbial and prepositional heads (Bogard & Company 1989, Girón 2004, Serradilla 2010, Pountain 2014).

In the case of *en caso que*, the introduction of *de* means that the construction cannot have been completely grammaticalized and non-compositional, otherwise *de* would not have entered it. It is, however, difficult to determine whether *caso* was analyzed as a noun, and *de* was added as part of the new nominal complement clause construction and its [N *de que*] schema, or if *en caso* was recognized as an adverbi(ial) expression (the [Adv_*de que*] schema) or, even, as a compound preposition *en caso de* ([preposition + *que*] schema). Given the existence of previous cases of *en el caso que*, however, the nominal complement clause construction seems like the most obvious alternative.

Following the figures in Table 4 (see Section 4.1, above), it is natural to compare the variants *en caso que* and *en caso de que* from 1500 to 1799, when they

dominate the construction. In fact, in my sample *en caso que* (198 cases) and *en caso de que* (116 cases), together make up over half of the 536 analyzed cases. But in between 1500 and 1799 the situation is even more extreme: the combined counts of *en caso que* and *en caso de que* amount to 82% (50/61), 78% (46/59), and 63% (69/110), respectively, of all the analyzed cases. Examples (10) and (11) are representative of the late 16th and early 17th century:

(10) pero que **en caso que** Su Majestad se determinase de hacer este viaje, le suplicaban mirase mucho de qué personas se confiaba y fiaba,
(Fray Prudencio de Sandoval, *Historia de la vida y hechos del Emperador Carlos V*, 1604–1618, CORDE)
'but in case His Majesty decided to make this voyage, they asked him to watch out carefully in which people he confided and trusted.'

(11) Dios mío, **en caso de que** algún loco se os quiera igualar no conociendo la infinita superioridad que tenéis á todo lo criado, es poco confesarla yo con mi boca.
(Fray Alonso de Cabrera, *Consideraciones sobre los Evangelios de los domingos de Adviento*, 1598)
'My Lord, in case some madman wanted to be equal to You without knowing how infinitely superior you are to everything [that has been] created, it is a small thing that I confess it by my mouth.'

As the examples show, the usage contexts for both variants are quite similar except for one detail: *en caso que* is more frequently combined with the past subjunctive (*se determinase* 'decide' in (10)), while *en caso de que* is more often followed by a present subjunctive such as *reciba* 'receive' in (11). Otherwise, both variants prefer the canonical [if A, B] type of conditional sentences and express possible conditions: it is possible that 'His Majesty might decide something', and also that 'some madman might want to do something'.

However, as Table 10 shows, although there is a higher usage percentage for *en caso que* with the past subjunctive (Example (10)) as compared to *en caso de que*, which prefers the present subjunctive (Example (11)), the difference in the verb forms used in the protasis is not significant. Even less significant differences are there in the types (Table 9) and kinds (Table 11) of conditionality associated with the *en caso que* and *en caso de que* variants.

In sum, the non-significant p-values in Tables 9 to 11 indicate that the introduction of *de* into the *en caso que* construction does not imply any real change in meaning. However, this introduction appears to be due first and foremost to the availability of the [prep_*que*] variable. In this context, the application of this variable is translatable into constructional terms as the activation of the [N *de que*] schema of the nominal complement clause construction (cf. Granvik 2015),

Table 9. Distribution of *en caso que* and *en caso de que* in different conditional sentences types (1500–1799)

	en caso que	*en caso de que*	**Sum**
if A, B	56 / 64%	50 / 63%	106 / 63 %
B, if A	22 / 25%	26 / 33%	48 / 29%
if A	9 / 11%	4 / 4%	13 / 8%
Sum	87 / 100%	80 / 100%	167 / 100%

X-squared = 2.3067, df = 2, p-value = 0.3156 *X-squared* = 2.3067, *df* = 2, *p-value* = 0.3156

Table 10. Distribution of different verb forms in the protasis after *en caso que* and *en caso de que* (1500–1799)

	en caso que	*en caso de que*	**Sum**
past subj	45 / 52%	29 / 38%	74 / 46%
pres subj	36 / 41%	45 / 58%	81 / 49%
fut subj	6 / 7%	3 / 4%	9 / 5%
Sum	87 / 100%	77 / 100%	164 / 100%

* *X-squared* = 4.8678, *df* = 2, *p-value* = 0.08769[19]

Table 11. Different kinds of conditional readings for the variants *en caso que* and *en caso de que* (1500–1799)

	en caso que	*en caso de que*	**Sum**
possible	85 / 98%	76 / 95%	161 / 96%
improbable	1 / 1%	4 / 5%	5 / 3%%
impossible	1 / 1%	0 / 0%	1 / 1%
Sum	87 / 100%	80 / 100%	167 / 100%

* *X-squared* = 3.015, *df* = 2, *p-value* = 0.2215.

which in the 16th century enters into competition with the earlier [N *que*] schema. This change in the nominal complement clause construction schema has been described as an analogical change due to functional motivations, whereby clausal complements came to be treated on a par with nominal and infinitival complements: N *de* N, N *de* INF vs. N *que* > N *de* N, N *de* INF, N *de que* (Bogard & Company, 1989; Barra, 2002; Serradilla, 2010; Granvik, 2017). From this perspec-

19. 19. As above, see footnote 18, the asterisk indicates that the expected value of over 20 percent of the cells was below 5, making the results of the X^2 test unreliable.

tive it should come as no surprise that little changes in meaning and usage are encountered in the data.

There is, a final noteworthy thing about Table 10, which becomes evident when comparing it to Table 7: there are no examples of the present indicative form. This might be taken as an indication that the CCCxn as a whole has moved one step further down the grammaticalization cline, or has entered a new level of schematicity, joining the paradigm of subordinating conjunctions governing the subjunctive mood. This is the system that has been active in Spanish since the 16th century, as indicated by expressions such as *a fin (de) que* 'in order that', *a condición (de) que* 'with the condition that', *con tal (de) que* 'with such that = if', which all govern the subjunctive. Actually accounting for this change, however, lies beyond the scope of this paper.

4.2.3 *Enter the determiner:* En caso de que *vs.* en el caso de que *from 1700 to 1899*

The last contrasting pair to be analyzed involves an element which is originally attested already in the medieval period, but its use only becomes numerically relevant in the 18th century. The alternation concerns the introduction of the determiner *el*, which completes the full variant of the CCCxn, *en el caso (de) que*. From a construction grammar perspective, the addition of *el* must, of course, be seen the opposite way, i.e. as the fact that the noun *caso* is inserted into the nominal complement clause schema [det N *que*+complement clause] (cf. Schmid, 2000; Granvik, 2015). This way, the determiner schema [det+N] can apply to *caso* as to any other noun which enters the nominal complement clause construction.[20]

The presence of *el* enhances the analyzability (or compositionality) of the complex conjunction, making *caso* individually more salient and referential. As such, following this step the conjunction as a whole seems less prone to be considered a chunk as compared to *en caso de que* or the other forms. An indication of the higher degree of analyzability of this construction format is that the variant *en el caso de que* also allows for the introduction of adjectival modifiers, such as *en el caso improbable de que[…]* 'in the unlikely case that[…]'.[21]

---

20. Note that directly using the determiner schema would be inaccurate since in the conditional *caso* construction only the definite article is possible, not the indefinite one. This restriction is discussed by Rostila (2004, § 4.2), but is possibly resolved for *en caso que* by relating this micro-construction to the nominal complement clause construction (N (*de*) *que*), which also does not usually permit the use of the indefinite article (cf. Leonetti, 1993, 1999).

21. Due to the way the analyzed sample was extracted, i.e. using the search strings *caso que* and *caso de que*, these cases could not be considered in this paper. Since CORDE is an unannotated corpus, inserting an adjective or modifier before and after *caso* is not an alternative.

Similarly to the two earlier subsections, the introduction of the determiner *el* in the construction does not mark a distinction between the two variants under analysis here. Instead, as the figures in Tables 12 to 14 show, in the period under investigation, there are very small differences between *en caso de que* and *en el caso de que* along the three dimensions considered: conditional sentence type, verb forms in the protasis and kind of conditional meaning.

Table 12. Distribution of *en caso de que* and *en el caso de que* in different conditional sentences types (1700–1899)

	en caso de que	*en el caso de que*	**Sum**
if A, B	42	44	86
B, if A	18	32	50
if A	2	2	4
Sum	62	78	140

* *X-squared* = 2.1662, *df* = 2, *p-value* = 0.3385

Table 13. Distribution of different verb forms in the protasis after *en caso de que* and *en el caso de que* (1700–1899)

	en caso de que	*en el caso de que*	**Sum**
past subj	25	27	52
pres subj	33	44	77
fut subj	0	4	4
Sum	58	75	133

* *X-squared* = 3.5331, *df* = 2, *p-value* = 0.1709

Table 14. Different kinds of conditional readings for the variants *en caso de que* and *en el caso de que* (1700–1899)

	en caso de que	*en el caso de que*	**Sum**
possible	60	68	128
improbable	2	4	6
impossible	0	2	2
Sum	62	74	136

* *X-squared* = 2.1244, *df* = 2, *p-value* = 0.3457

All in all, the findings regarding the three main changes in the constructional format (eliminating *en* to give *caso que*, adding *de* to give *en caso de que*, and adding *el* to give *en el caso de que*) can be taken to indicate that the successive addition of elements to the CCCxn has meant that the construction as a whole is

not stored as a single unit, or a chunk –although this possibly did happen with *caso que* in the earlier centuries–, but rather is analytically accessed and constantly available for further elaboration. This has interesting consequences for the two theoretical models I have attempted to apply in order to account for the development of the CCCxn.

4.3 Grammaticalization or constructionalization/constructional change?

In this final section of the analysis, I will address two main questions. I will begin by considering the implications that the formal variation has for the grammaticalization account of the CCCxn, and continue by assessing the same question with regards to T&T's (2013) model of constructionalization and constructional change. Despite the title to this subsection, I do not believe that opposing the notions is really fruitful. Instead, I will emphasize that both approaches are important for understanding the development of the conditional *caso* construction and its six variants.

The first question which arises is: does the development of the CCCxn count as grammaticalization? As stated earlier (see Section 4.1), out of Heine's (2003, p. 579) and Heine et al.'s (2016, p. 151) four criteria for grammaticalization, i) desemanticization; ii) context extension; iii) decategorialization and loss in morphosyntactic properties; and iv) erosion and loss in phonetic substance, I believe at least three are involved in the development of conditional *caso*. There is clearly a change in meaning from the lexical towards the more grammatical in the rise of the conditional meaning. This change in meaning is accompanied by a change in the range of uses (i.e. host-class extension, see Himmelmann, 2004): nominal *caso* is used as a noun, when combined with the preposition *en* 'in' it can be used to introduce a situation, instance or case. In this adverbial function, *caso* is compatible with a hypothetical interpretation, and after adding *que* the resulting conjunction establishes a (conditional) relationship between (two) propositions. The change also involves a loss of morphosyntactic properties of the noun *caso*, especially in the more fixed, determiner-less variants, such as *caso (de) que* and *en caso (de) que*, where it loses most of its referentiality.

However, these four criteria have little to say about the formal variation. Furthermore, from a grammaticalization perspective the formal changes to the conjunction pose a potential problem, since the addition of elements to a supposedly grammaticalized expression means that its main component, in this case *caso*, must be accessible to (sub)conscious analysis and recognizable as a noun.

From the perspective of this study, the grammaticalization model can be considered an accurate tool for describing the development of the grammatical *en caso que* and *caso que* construction from the starting point of the abstract noun

caso. If, as happened in Portuguese and Italian, the development had gone further and the compound conjunction *caso que* had lost also its final element, turning *caso* into a conjunction in and of itself, this change could also have been accounted for by the grammaticalization model (cf. Rosenkvist, 2004 for Swedish *ifall*). As the following quotes by Heine et al. (2016) and T&T (2013) underline, grammaticalization is at its best when accounting for the development of grammatical formants: [grammaticalization] "accounts for regularities in the evolution of grammatical (or functional) categories," (Heine et al., 2016, p. 168); and "the findings of grammaticalization[…] provide the evidence for the micro-changes that lead to […] grammatical constructionalization" (T&T, 2013, p. 148).

However, the grammaticalization model is less well equipped to account for the fact that in the development of conditional *caso*, there is not a only unidirectional move from *caso* (noun) to *caso que* (conjunction), but also a 'survival' and even a 'revival' of the original complex expression, *en caso que*. This brings us to the second question, the constructionalization of *caso que*. When *en caso que* is analyzed as a compound expression with a nominal base, and the appropriate variant of the [prep_*que*] variable, e.g. *de que*, is added to the base, we are no longer talking about a grammaticalized unit or a chunk. Instead, there is a constructional schema which is open and accessible to the speakers who can insert the CCCxn base into other productive schemas of different levels. That is, we can talk about the constructionalization of conditional *caso*.

In a similar fashion, when the determiner *el* was introduced in the conditional conjunction in the 14th century, the nominal complement clause construction had the basic schematic format [(det)+N+*que*+clause]. *En caso que* was thus found to be a suitable member of the nominal complement clause construction, and sanctioned by this schema, whereby the determiner *el* could be used. According to the data in Table 4 (Section 4.1, above), there was a second round of integration into the nominal complement clause construction in the 18th century. By this time this constructional schema had also become associated with the [prep_*que*] schema, so that it had the format [(det)+N+*de*+*que*+clause], yielding the form *en (el) caso de que*.

In Figure 1, I have tried to illustrate the constructional network of the Conditional *Caso* Construction, including the six formal variants and incorporating some higher level constructional schemas which are tightly connected to the CCCxn. These constructional schemas include, on the one hand, the paradigm of Spanish complex conjunctions (see Table 2, above), and the nominal complement clause construction, which are situated above the conditional *caso* schema (in the middle). On a higher level still, affecting these two mid-level schemas as well as the lower-level micro-constructions, I have included both the [prep_*que*] schema

and the [det+N] schema, in order to account for the introduction of *de* and *el* into the conjunction.

Figure 1 is, of course, a two-dimensional simplification of the real network structure of which the CCCxn is a part. However, I believe it helps clarify some important points. First, it includes important inheritance links from one constructional schema to another. These are shown by dotted straight and dashed curved lines. The dotted straight lines show the inheritance links from the [prep_*que*] schema, while the dashed curved lines indicate inheritance links from the [det+N] schema. The continuous lines indicate a schema–subschema or subschema–micro-construction relationship. For example, the CCCxn is shown as essentially divided in two, with the *caso que* and *en caso que* variants representing two different nodes. This split is due to the fact that only in the *en caso que* node is the determiner added, i.e. only this part of the constructional schema is affected by the inheritance links from the nominal complement clause construction and, from an even higher level, the links from the [det+N] schema (dashed curved lines). The top left hand side of Figure 1 includes the [prep_*que*] schema, which applies on both the complex conjunction construction (affecting the *caso que* micro-schema) and the nominal complement clause construction (affecting the *en caso que* node) (dotted straight lines).

Second, the conditional *caso* construction is sanctioned by two constructional schemas, the complex conjunction schema and the nominal complement clause schema. These two schemas are also related, in the sense that there are many complex conjunctions (and compound prepositions) the center of which is a noun (e.g. *con la condición de que* 'with the condition [of] that', *a pesar de que* 'in spite [of] that', *en el sentido de que* 'in the sense [of] that'. However, the precise details of this relationship cannot be discussed here.[22]

Third, Figure 1 includes six substantial micro-constructions, which correspond to the six formal variants. Since they have been assigned their own nodes, an orthodox interpretation of the constructional network structure has it that they are different constructions, i.e. they correspond to different form–meaning pairings. According to Goldberg's (1995) Principle-of-No-Synonymy, for example, a difference in form implies a difference in meaning. However, the analysis presented in the previous sections casts some doubts upon whether this is actually the case. At least according to the three criteria applied in this study, there does not seem to be any significant semantic and functional differences between the three

22. As suggested to me by a reviewer, Van de Velde's (2014) horizontal links might be a useful means of describing the relationship between the nominal complement clause and the complex conjunction schemas, but space and time limitations disallow me from developing this question further.

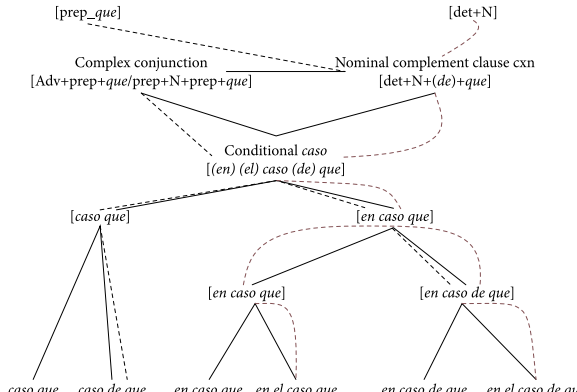


Figure 1. Proposed network structure for the conditional *caso* construction. The square brackets indicate constructional schemas, and the italicized expressions substantial constructional variants

main pairs of the six variants. The six micro-construction might thus be considered allostructions in Cappelle's (2006) sense, a point I will return to below.

From a diachronic perspective, following T&T's (2013) distinction between constructionalization and constructional changes, the only case of constructionalization would be the creation of the constructional caso construction schema. Given that *en caso que* is both the earliest and the most frequently attested variant overall (at least in my data), it seems reasonable to identify it as the centre of the outcome of the constructionalization process. However, this constructionalization process has been followed by a series of constructional changes. These have not only affected the status of *en caso que* in particular, but modeled the whole construction so that it does not simply consist of a single form with different related uses, but is a family of forms that have a corresponding family of uses. That is to say, instead of being a polysemous construction, which is, in a sense, the expected outcome of constructionalization and subsequent constructional changes, the CCCxn resembles a family of synonymous formally-distinct micro-constructions (or allostructions).[23]

Apart from offering an interesting counter-example to Goldberg's Principle-of-No-Synonymy, the six variants of the CCCxn also seem to underplay the "competing motivations" involved in Goldberg's (1995, pp. 67–68) Principle of Maximized Expressive Power and Principle of Maximized Economy (discussed in T&T (2013, pp. 146–147)): the Spanish conditional *caso* has developed six formal variants, which all seem mutually interchangeable.

The important point here, though, is not the fact that the six variants may go counter to certain principles, but the theoretical status of the six variants. Are they different constructions or not? Construction grammar does not seem to offer any immediate solution to this problem. Since a difference in form usually implies a difference in meaning, postulating that distinct nodes in a network correspond to different constructions, and that different constructions differ in meaning as well as in form seems natural enough. However, despite looking quite thoroughly into the interplay between form and meaning in the CCCxn, this does not seem to be the case.

From a theoretical perspective, the most natural explanation of this state-of-affairs is to say that there must be some difference in meaning or at least in the usage context of the six variants. And indeed there probably are: the *de*-less vari-

23. This means that our construction grammar approach is now facing the same problem as the grammaticalization model did earlier: the outcome of change is not what is theoretically expected, i.e. […] the expected result of a grammaticalization process is a new grammatical formant with a new grammatical function, not the coexistence of variants with essentially the same function.

ants are definitely not preferred in formal and written registers of present-day Spanish, and their usage frequency is very limited as well (see Tables 4 and 5, above). But there are three variants which make up two alternation pairs, i.e. *en caso que* vs. *en caso de que* on the one hand, and *en caso de que* and *en el caso de que*, on the other. These pairs do not seem amenable to an easy solution such as the one just offered, but instead seem very much interchangeable. Considering the kind of elements that mark the formal difference, the determiner *el* and the most abstract of Spanish prepositions *de*, a low degree of semantic contrast and contextual interchangeability is, in fact, to be expected.

This brings us back to the question of the theoretical status of these variants. How should these near-synonyms be treated from a construction grammar perspective? What status can they have as members of the constructional network? My initial solution was to treat them as distinct micro-constructions and include them in the constructional network (as pictured in Figure 1). But I acknowledge that this is not unproblematic, especially since the differences in usage and meaning are so slight.

Following Cappelle (2006) (see also Perek [2015, Chapter 6]), a more appealing solution might be to consider the six micro-constructions as allostructions of the conditional *caso* construction, that is, "as variant structural realizations of a construction that is left partially underspecified" (Cappelle, 2006, p. 19).[24] This alternative has the advantage of allowing the incorporation of synonymous forms into the constructional network. Viewing alternations as (horizontal) links between allostructions, as Cappelle (2006) does in his study on the English transitive verb-particle construction, is an obvious option also for the formal variants of the conditional *caso* construction. As the denotation of the CCCxn in Figure 1 shows, the presence of (*en*), (*el*), and (*de*) is not obligatory, so these can be left unspecified at the level of the CCCxn schema (cf. Figures 1 and 2 in Cappelle [2006, pp. 18, 24]).

The discussion of the exact level(s) where the horizontal links should be introduced, however, is an open one, which is why they have not been made explicit in Figure 1. In order to determine the actual degree of interchangeability and alternation between the variants, more empirical evidence is needed, but this must be left for future research.

24. Sadly, space and time limitations make it impossible for me to fully develop Cappelle's (2006) notion of allostructions in this paper.

5. Conclusions

In this paper I have applied two models of linguistic change on an empirical case study of the conditional *caso* construction in Spanish. The aim has been to compare how the grammaticalization and constructionalization models fare in accounting for the creation of a new grammatical meaning and the subsequent formal changes in the conditional *caso* construction. I have tried to show that both models are useful and important, but due partly to their different interests or focus (cf. Heine et al., 2016), partly to the phenomenon at hand, they are so with regards to different aspects of the change. From the perspective of conditional *caso* in particular, I do not see a reason for considering them mutually exclusive.

When it comes to the development of the conditional reading of *caso*, the grammaticalization model provides an accurate toolkit for describing the change. However, when moving beyond the directional change from noun to gram, the grammaticalization model has less to say about the formal changes. In fact, the coexistence and reactivation of forms that according to the model are 'earlier' or higher up the cline do not fit easily into the grammaticalization view on change.

Instead, when addressing these aspects, the constructionalization and constructional change model by Traugott & Trousdale (2013) provides a better toolbox. Their concepts of schematicity, productivity and compositionality are conveniently applied when the focus lies not on the micro-level development of a gram, but on the broader context of this change. By relating the conditional *caso* construction to other constructional schemas, such as the complex conjunction, the nominal complement clause and the [prep_*que*] variable, the changes to the formal architecture of the CCCxn become much clearer.

The network model of Construction Grammar undoubtedly provides a powerful tool for dealing with the formal changes to a construction, allowing one to relate the different forms to higher and lower level schemas. However, the fact that my analysis shows little semantic and functional (or distributional) differences between the six formal variants means that the final descriptive step is not entirely clear: what is the status of the micro-constructions? Here an extension of the Construction Grammar view on grammatical structure provides a convenient and psychologically plausible solution in the form of Cappelle's (2006) notion of allostructions, since his model integrates synonymy into the network model of Construction grammar. Considering the CCCxn as "partially underspecified" and the six variants as "subforms" (Cappelle, 2006, p. 25) or allostructions is definitely an option.

The allostruction model does not mean, however, that drawing the line between allostructions and different micro-level constructions is a straightforward task. As the present paper has shown, with regards to conditional *caso* more

emipirical work is definitely in place in order to answer this question. On a more general level, it is not completely clear what kind of evidence is needed to prove the existence of synonymy links, especially when grammatical elements such as determiners and abstract prepositions are concerned. Thus, there is certainly a need for future studies where the synonymous status of two different forms is not explained away, but rather integrated in the descriptive model.

References

Barðdal, J., Smirnova, E., Sommerer, L., & Gildea, S. (Eds.). (2015). *Diachronic Construction Grammar*. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.18

Barra Jover, M. (2002). *Propiedades léxicas y evolución sintáctica. El desarrollo de los mecanismos de subordinación en español*. A Coruña: Toxosoutos.

Bartens, A., & Granvik, A. (2012). Gramaticalización y lexicalización en la formación de locuciones preposicionales en español e italiano. In E. Montero Cartelle (Ed.), *Actas del VIII Congreso Internacional de Historia de la Lengua Española* (Universidade de Santiago de Compostela, 14–18 September 2009), Volume 2 (pp. 2029–2046). Santiago de Compostela: Meubook.

Bartol Hernández, J. A. (1986). *Oraciones consecutives y concesivas en las Siete Partidas*. Acta Salmanticensia: Studia Philologica Salmanticensia.

Bogard, S., & Company Company, C. (1989). Estructura y evolución de las oraciones completivas de sustantivo en el español. *Romance Philology*, XLIII(2), 258–273.

Brinton, L., & Traugott, E. C. (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615962

Cano Aguilar, R. (2010). Aspectos discursivos en la historia de los períodos oracionales complejos en castellano medieval. In R. M. Castañer Martín, & V. Lagüens Gracia (Eds.), *De moneda nunca usada: Estudios dedicados a José Mª Enguita Utrilla* (pp. 149–162). Zaragoza: Insituto Fernando El Católico, CSIC.

Cappelle, B. (2006). Particle placement and the case for 'allostructions'. *Constructions*, SV1(7), 1–28.

Cifuentes Honrubia, J. L. (2003). *Locuciones prepositivas – sobre la gramaticalización preposicional en español*. Alicante: Universidad de Alicante.

Codita, V. (2013). Hápax y otras construcciones prepositivas de frecuencia escasa en la *Primera Partida* . *Revista de Historia de la Lengua Española*, 8, 63–92.

Collins Spanish Dictionary. Sixth edition. (2000). Glasgow: Harper Collins.

[CORDE] = Real Academia Española: Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*. <http://www.rae.es> [March 2016]

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Devís Márquez, P. P. (1994). El concepto de subordinación. Criterios para la clasificación de las denominadas oraciones subordinadas en español. *Contextos*, XII(23–24), 71–106.

[DRAE] = Real Academia Española. (2014). *Diccionario de la lengua española*. 23rd edition. http://dle.rae.es/.

Fagard, B., & De Mulder, W. (2007). La formation des prépositions complexes: grammaticalisation ou lexicalisation? *Langue française*, 156(4), 9–29.

Garrido Sepúlveda, C. (2015). La restructuración de las oraciones condicionales desde el latín al español medieval: hacia un estado de la cuestión. *Onomázein*, 31, 187–204. https://doi.org/10.7764/onomazein.31.12

Garrido Vílchez, G. B. (2008). *Las gramáticas de la Real Academia Española: teoría grammatical, sintaxis y subordinación (1854–1924)*. PhD dissertation. Salamanca: Universidad de Salamanca.

Girón Alconchel, J. L. (2004). Cambios gramaticales en los Siglos de Oro. In R. Cano Aguilar (Ed.), *Historia de la lengua española* (pp. 859–893). Barcelona: Ariel.

Girón Alconchel, J. L. (2008). Lexicalización y gramaticalización en la creación de marcadores del discurso… y de otras palabras. In E. Stark et al. (Eds.), *Romanische Syntax im Wandel* (pp. 363–385). Tübingen: Gunter Narr Verlag.

Goldberg, A. (1995). *Constructions. A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, A. (2006). *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Granvik, A. (2013). Adverbios, preposiciones y conjunciones subordinantes: el papel de *de* y *que* en el establecimiento de categorías gramaticales contrastivas en español y portugués. In E. Havu, & I. Hyvärinen (Eds.), *Comparing and contrasting syntactic structures. From dependency to quasi-subordination* (pp. 67–96). Helsinki: Société Néophilologique.

Granvik, A. (2014). Hablando *de*, sobre y *acerca de* la gramaticalización y la lexicalización: panorama diacrónico de las relaciones entre preposiciones y locuciones prepositivas dentro del campo semántico de TEMA/ASUNTO . In J. L. Girón Alconchel, & D. M. Sáez Rivera (Eds.), *Procesos de gramaticalización en la historia del español* (pp. 77–117). Madrid/Frankfurt Am Main: Iberoamericana/Vervuert.

Granvik, A. (2015). Oraciones completivas de sustantivo: un análisis contrastivo entre español y portugués. *Verba*, 42, 347–401. https://doi.org/10.15304/verba.42.1856

Granvik, A. (2017). Oraciones completivas de sustantivo en español y portugués: ¿infinitivo u oración finita? *Cuadernos de Lingüística de El Colegio de México*, 5, 103–180. https://doi.org/10.24201/clecm.v4i1.54

Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 83–101). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Heine, B. (2003). Grammaticalization. In B. D. Joseph, & R. D. Janda (Eds.), *Handbook of Historical Linguistics* (pp. 575–601). Oxford: Blackwell. https://doi.org/10.1002/9780470756393.ch18

Heine, B., Narrog, H., & Long, H. (2016). Constructional change vs. grammaticalization. *Studies in Language*, 40(1), 137–175. https://doi.org/10.1075/sl.40.1.05hei

Himmelmann, N. P. (2004). Lexicalization and grammaticalization. Opposite or orthogonal? In W. Bisang, N. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization? A Look from its Fringes and its Components* (pp. 21–42). Berlin: Mouton de Gruyter.

Hoffmann, S. (2004). Are low-frequency complex prepositions grammaticalized? On the limits of corpus data – and the importance of intuition. In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to Grammaticalization in English* (pp. 171–210). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.13.09hof

Hoffmann, S., & Trousdale, G. (2013). Introduction. In S. Hoffmann, & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 1–15). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195396683.001.0001

Ibba, D. (2007). Oraciones concesivas y gramaticalización: el caso de *aunque* y *maguer (que)* . *Interlingüística*, 17, 493–502.

Julián Mariscal, O. (2009). La condicionalidad de las locuciones sin *si* frente a las pseudocondicionales. Estudio dieciochesco. *Interlingüística*, 18, 614–624.

Lehmann, C. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile*, 20, 303–318.

Lehmann, C. (2002). New Reflections on Grammaticalization and Lexicalization. In I. Wischer, & G. Diewald (Eds), *New Reflections on Grammaticalization* (pp. 1–18). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.03leh

Lehmann, C. (2016). Complex spatial prepositions from Latin to Spanish. Paper presented as part of the Workshop on Emergence and Evolution of Complex Adpositions in European Languages at the 49 Meeting of the Societas Linguistica Europaea, 31 August – 3 September 2016, Naples. http://www.christianlehmann.eu/publ/lehmann_complex_prepositions.pdf.

Leonetti, M. (1993). Dos tipos de completivas en sintagmas nominales. *Lingüística (ALFAL)*, 5, 5–40.

Leonetti, M. (1999). La subordinación sustantiva: las subordinadas enunciativas en los complementos nominales. In I. Bosque, & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 2083–2104). Madrid: Espasa Calpe.

Marcos Marín, F. (1979). Observaciones sobre las construcciones condicinales en la historia de la lengua española. *Nueva Revista de Filología Hispánica*, XXVIII, 86–105.

Meillet, A. (1975[1912]). L'évolution des formes grammaticales. In *Linguistique historique et linguistique générale* (pp. 130–148). Paris: Société de Linguistique de Paris.

Meillet, A. (1975[1916]). Le renouvellement des conjonctions. In *Linguistique historique et Linguistique générale* (pp. 159–174). Paris: Société de Linguistique de Paris.

Mondéjar, J. (1966). La expresión de la condicionalidad en español. *Revista de Filología Española*, 49, 229–254. https://doi.org/10.3989/rfe.1966.v49.i1/4.875

[NGLE] = Real Academia y Española y Asociación de Academias de la Lengua Española. (2009). *Nueva Gramática de la Lengua Española*. Madrid: Espasa.

Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language*, 14(2), 177–202. https://doi.org/10.1075/fol.14.2.04noe

Pavón Lucero, M. V. (1999). Clases de partículas: preposición, conjunción y adverbio. In I. Bosque, & V. Demonte (Eds.), *Gramática descriptive de la lengua española* (pp. 565–655). Madrid: Espasa Calpe.

Penny, R. (2002). *A History of the Spanish Language*. Second edition. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511992827

Perek, F. (2015). *Argument structure in usage-based construction grammar: experimental and corpus-based perspectives*. Amsterdam: John Benjamins.

Pountain, C. (2014). Preposición + *que* en español. *Cuadernos de Lingüística de El Colegio de México*, 2, 9–54.

Rosenkvist, H. (2004). *The Emergence of Conditional Subordinators in Swedish – a Study in Grammaticalization*. Dissertation. Lund: Department of Scandinavian Languages, Lund University.

Rostila, J. (2004). Lexicalization as a way to grammaticalization. Proceedings of the 20th Scandinavian Conference of Linguistics, Helsinki, January 7–9, 2004.
Rostila, J. (2006). Storage as a way to grammaticalization. *Constructions*, 1.
Sansò, A., & Mauri, C. (2014). Pathways to conditionality: two case studies from Italian. *Archivio Glottologico Italiano*, XCIX(1), 97–121.
Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110808704
Serradilla Castaño, A. (2010). La subordinación completiva en español antiguo: continuidad y ruptura de los modelos latinos. In M. Castillo Lluch, & M. López Izquierdo (Eds.), *Modelos latinos en la Castilla medieval* (pp. 145–158). Madrid/Frankfurt Am Main: Iberoamericana/Vervuert.
Trousdale, G., & Norde, M. (2013). Degrammaticalization and constructionalization: two case studies. *Language Sciences*, 36, 32–46. https://doi.org/10.1016/j.langsci.2012.03.018
Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001
Van de Velde, F. (2014). Degeneracy: The maintenance of constructional networks. In R. Boogaart, T. Colleman, & G. Rutten (Eds.), *Extending the Scope of Construction Grammar* (pp. 141–180). Berlin: De Gruyter.

Constructionalization areas

The case of negation in Manchu

Andreas Hölzl
Ludwig-Maximilians-Universität München, Germany

Heine & Kuteva (2005) proposed the concept of a *grammaticalization area*. Given the recent appearance of the concept of *constructionalization*, (Traugott & Trousdale, 2013) it seems reasonable to ask whether there are *constructionalization areas* as well. The paper exemplifies this concept with the help of the Tungusic language Manchu and some surrounding languages. Manchu does not have the common Tungusic negative verb *e-*, but, like Mongolian, employs the negative existential *akū* instead (e.g., Ikegami, 1999 [1979]). Manchu and Mongolian share not only this development, but also an underlying *interlingual constructional network* (Höder, 2012). The paper proposes the name *areal construction grammar* as a label for those aspects of construction grammar that are concerned with the study of language contact.

Keywords: Tungusic, Manchu, Mongolic, Khalkha Mongolian, negation, negative existential cycle, shared constructionalization, constructionalization area, interlingual constructional network, areal construction grammar

1. Introduction[1]

Construction grammar (CxG) is a growing field of inquiry that is being extended to new areas such as historical linguistics, which has led to so-called *diachronic construction grammar* (e.g., Noël, 2007). This paper argues that it can also serve as a basis for areal linguistics in what may be called *areal construction grammar*. In this sense, this paper aims at contributing to the schematicity of CxG by applying it to yet another subfield of linguistics (cf. Barðdal & Gildea, 2015, p. 42). For this

1. I want to thank Elizabeth Traugott, Tom Payne, and the anonymous reviewers for their comments on earlier versions of this paper.

https://doi.org/10.1075/cal.21.c9

purpose, it integrates certain aspects of previous research on language contact and shared grammaticalization (e.g., Heine & Kuteva, 2010; Aikhenvald, 2013).

The basic units of CxG are constructions, conventionalized and entrenched pairings of form and meaning. They come in different degrees of analyzability (from analyzable to non-analyzable), complexity (from atomic to complex) and schematicity (from schematic to specific) and may be integrated with each other more or less freely (e.g., Langacker, 2008). Complex constructions have an "external" holistic meaning that is different from the sum of the meaning of their "internal" constituents, which nevertheless have their own properties (Fried, 2013). Constructions do not stand on their own but are part of larger networks of related constructions and always occur in actual usage (as so-called constructs). The meaning of constructions ranges from contentful to procedural (rather than from lexical to grammatical) (Trousdale, 2014, p. 559). Diachronic construction grammar recognizes an additional diachronic dimension and makes a distinction between constructional change (changes within an existing construction) and constructionalization (the creation of a new form-meaning pair of any kind) (Traugott & Trousdale, 2013). It has been suggested that constructions may be shared by different languages (e.g., Höder, 2012; Hilpert & Östman, 2014), and this paper argues that languages may in addition also share instances of constructionalization.

This paper is about similarities between languages, especially similarities of constructions, constructional networks, and constructionalization. But languages may be similar for several reasons, including universals or strong tendencies, chance resemblance, parallel development, genetic inheritance, and contact (e.g., Aikhenvald, 2013). The focus here lies on language contact. However, it is not always easy to distinguish clearly between all these different possibilities and in most cases several causes apply simultaneously (Thomason, 2010). It is no news that constructions (form and/or meaning) may be adopted by one language from another. But language contact may also lead to shared *changes* in form, meaning, or both. This paper investigates such changes in a number of languages in Northeast Asia with a focus on Tungusic, less so on Mongolic, and least on some surrounding languages. The type of change investigated is the so-called *negative existential cycle* (Croft, 1991; Veselinova, 2014, 2015, 2016), which is a cyclical change based on six different types or stages (Figure 1).

> In *type A*, the SN [standard negation] strategy is used to negate both verbal and existential sentences. In *type A~B*, there is a special existential negator that is limited to particular contexts; for example, this type includes languages in which the special existential negator is restricted to the present tense. *Type B* comprises languages in which existential predications are negated by a special strategy only.

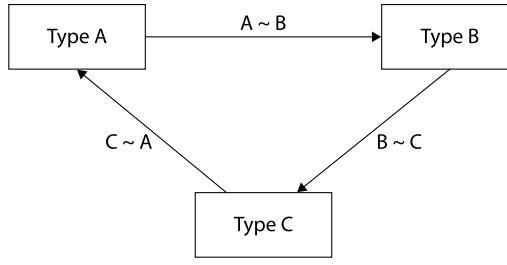


Figure 1. The negative existential cycle (cf. Veselinova, 2014, 2015, 2016)

> *Type B~C* covers languages in which the special existential negator is also used to negate verbs in certain contexts/grammatical categories […]. In languages classified as *type C*, the existential negator is regularly used to negate verbs; however, the strategy used to negate existential predication still differs from SN in one or more respects. […] The final stage in the completion of the cycle is represented by *type C~A*, whereby the negative existential negator has begun to be used with the affirmative existential to yield emphatic/pragmatically marked constructions.
>
> (Veselinova, 2014, pp. 1328f., my emphasis and brackets)

A majority of languages in Northeast Asia belong or belonged to type B. The Tungusic languages Evenki and Udihe, for instance, exhibit a negative verb *e-* '(to do) not' for standard negation, but have a different negator for negative existence (Evenki *aachin*, Udihe *anchi* 'there is/are no') (e.g., Nedjalkov, 1994; Hölzl, 2015). This paper investigates those languages that show a development from type B to type C and beyond. Of course, this is a simplified picture as the negators are always embedded into larger constructions.

In Heine & Kuteva's (2010, p. 97) terminology, the languages analyzed in this paper would perhaps qualify as a *grammaticalization area*, "a group of geographically contiguous languages that have undergone the same grammaticalization process as a result of language contact." Two languages cannot, of course, have exactly the *same* process but only an *analogous* one. The shared development is based on language specific constructions that may be similar to a certain degree, but is always based on an analogy between these languages. However, there usually are at least differences in the form side as well as in many details during the development. Other authors employ the term *shared grammaticalization* instead (e.g., Aikhenvald, 2013).

But CxG offers a broader approach to language change and language in general than does grammaticalization (e.g., Traugott & Trousdale, 2013). Consider the example of demonstratives (Diessel, 2006). They are traditionally considered part of grammar, and as expected they may be the source of "more grammatical" elements such as third person pronouns (e.g., Manchu *te.re*, Mongolian *te.r*

'that > 3SG'). But demonstratives, apart from reinforcement with lexical elements (e.g., Manchu *tu-ba-* 'that-place-'), do not usually originate in the lexicon, as would be expected for grammatical elements, although there are several counterexamples (Heine & Kuteva, p.c. 2017). Consequently, Diessel (2006, p. 482) concludes that they "must be kept separate from all other linguistic expressions" (except for interrogatives). In other words, they are located outside the continuum from lexicon to grammar that forms the very foundation of grammaticalization research. For grammaticalization they might constitute an anomaly, but for CxG they are simply a different class of constructions with a specific form and a specific procedural meaning (i.e. joint attention, Diessel, 2006). CxG also allows the existence of multi-modal constructions (e.g., Steen & Turner, 2013). Demonstratives are usually accompanied by co-speech gesture, especially indicating with fingers, head, lips, or eyes. Without an additional gesture, demonstratives are often ambiguous. Imagine the sentence *that woman over there* uttered during a crowded conference break. Without pointing out the specific person, the expression is non-referential. Grammaticalization has no means of capturing these facts. But this does not mean that some key insights from grammaticalization cannot be usefully integrated into the newly emerging historically oriented CxG. Loosely based on the approaches mentioned above, this paper argues for the existence of *shared constructionalization* and *constructionalization areas*. The latter term may be defined as a group of often geographically contiguous languages that share an analogous instance of constructionalization due to language contact (cf. Heine & Kuteva, 2010, p. 97) . The term *area* here has to be understood in a very loose way and is not intended to indicate any clear-cut boundaries. Manchu (Tungusic) and Mongolian (Mongolic) are especially good and well-known examples of such an area (e.g., Ikegami, 1999 [1979]), which is why most examples will be drawn from these languages. In principle, the model proposed here is still close to Heine & Kuteva's (2010, p. 86) *grammatical replication*, "a process whereby speakers create a new grammatical meaning or structure in language R on the model of language M by using the linguistic resources available in R", of which *contact-induced grammaticalization* is a subtype. However, the precondition for a shared constructionalization, and especially a similar result of the kind investigated here is a shared source construction that is part of a larger network of constructions. Furthermore, "it is the construction as a whole, and not the negative existential alone, that expands its domain of use" (Veselinova, 2014, p. 1351). The more similar the source constructions are, the more similar are the results. The development addressed here qualifies as an instance of "grammatical" constructionalization and is thus necessarily similar to traditional grammaticalization. But in order to emphasize the broader scope, some instances of "lexical" construction-

alization (or lexicalization) that go beyond the negative existential cycle will be given as well (Traugott & Trousdale, 2013).

Originally, the negative existential cycle was postulated on the basis of synchronic data (Croft, 1991), but Veselinova (2014, p. 1373) is certainly right in her assumption that "it is indispensable that such work be supplemented by as many historical-comparative studies as possible". In sum, this study addresses the following research questions: Which languages in Northeast Asia exhibit the negative existential cycle? Do they follow the negative existential cycle precisely as postulated by Croft? How can CxG be applied to areal linguistics? Can instances of constructionalization be shared by different languages? What are the similarities and differences among the developments in Manchu, Mongolian, and the surrounding languages? On what constructional networks are the developments based? When did the process take place?

This paper has five sections, including this Introduction. Section 2 gives an account of the negative existential cycle in some Tungusic languages. Sections 3 and 4 address Mongolic and surrounding languages with more or less similar developments, respectively. Section 5 gives some conclusions for the negative existential cycle and areal CxG. Figure 2 indicates the approximate location of the languages addressed within this paper.

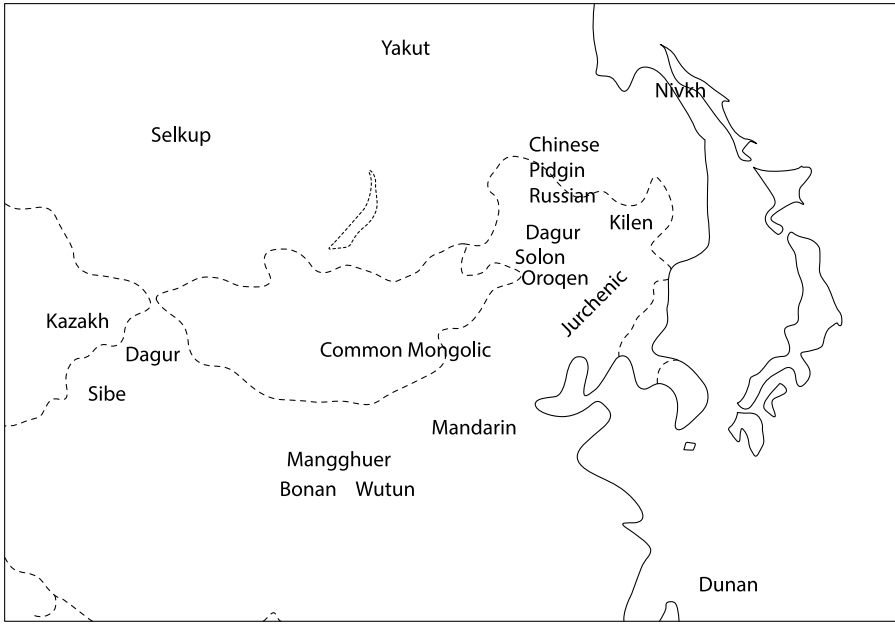


Figure 2. The approximate location of the languages surveyed. Common Mongolic, Kazakh, Yakut, and Mandarin have large geographical distributions not shown in detail

2. Tungusic

Tungusic languages are spoken in Siberia and northern China. They are all endangered, moribund, or already extinct and together have less than 100,000 speakers. They may be classified as in (1) (Janhunen, 2012b, p. 16).

(1) 1. Northern Tungusic
    1.1. Ewenic (†Arman, Even, Evenki, Negidal, Khamnigan Evenki, Oroqen, Solon)
    1.2. Udegheic (Udihe, Oroch)
  2. Southern Tungusic
    2.1. Nanaic (Nanai, Kili, Kilen, Ulcha, Uilta)
    2.2. Jurchenic (see 2)

Manchu is the best known and most prominent representative of the Jurchenic branch of Tungusic. The name *Manchu* was officially introduced as late as in 1635 as a cover term in order to replace older names such as *Jurchen*. But Jurchenic shows stronger internal variation than previous approaches assumed and in my current classification can be divided into three subgroups. Dates refer to the approximate times of the recordings.

(2) 1. Manchuic
    1.1. †Jurchen^{B} (about 1500?)
    1.2. Written Manchu (since about 1600)
    1.3. Manchu dialects (20th and 21st centuries)
      1.3.1. (†)Manchurian dialects
      1.3.2. Dzungarian dialect(s) = Sibe
  2. Balaic
    2.1. †Written Jurchen (12th–15th century)
    2.2. †Jurchen^{A} (about 1450–1500?)
    2.2. †Bala (20th century)
  3. Alchukaic
    3.1. †Alchuka (20th century)

Manchuic, the largest and best known subgroup, comprises Written Manchu and modern Manchurian dialects such as that in Aihui or Sanjiazi and Sibe in Dzungaria (northern Xinjiang). Manchu was the language of the Qing 清 dynasty (1636–1912), which conquered large parts of East and Central Asia and represents the last dynasty of imperial China. But the number of speakers has been declining in favor of Chinese since the 18th century. By now there are almost no speakers left in Manchuria and the last speakers of a Manchuic language can be found in Dzungaria at the Kazakh border. This declining group of perhaps still up

to 30,000 speakers (cf. Zikmundová, 2013, p. 10) belongs to the officially recognized Sibe minority. But the majority of the Sibe still live in Manchuria and have lost their language as well. Written Manchu with its script based on the Mongolian alphabet has existed from about 1600 and to a certain degree is still in use by the Sibe in Xinjiang. One only historically attested Manchuic vernacular was recorded about a hundred years or so before Written Manchu was created. This language will be called Jurchen[B] here and is only known through transcription with Chinese characters (Kane, 1989). Balaic has the modern member Bala (Mu Yejun, 1987) as well as one historically attested language called Jurchen[A], which had its own literary language and scripts (Kiyose, 1977). It was the language of the Jin 金 dynasty (1115–1234), which conquered what today is northern China. Bala may have a slightly different dialectal basis than Jurchen[A] and in addition shows influence from other Tungusic languages (e.g., Hölzl, 2015, p. 135, footnote 27, 2017). Alchukaic only has one modern member, which will simply be called Alchuka (Mu Yejun, 1986) and is not attested historically. Balaic, Alchukaic, and most of Manchuic are extinct. A detailed justification for this tentative classification goes beyond the scope of this paper (but see Mu Yejun, 1986, 1987; Kiyose, 2000). Some remarks concerning the special status of Alchuka will be made in this section.

There are no easily accessible large scale electronic corpora for Jurchenic languages that could be used to track the development over time. This rules out the possibility of statistical approaches, which will have to be included in future studies.

Tungusic languages usually have the typologically interesting phenomenon of conjugated negative verbs (e.g., Nedjalkov, 1994). In the Nanaic language Uilta this verb has the form *e-* and takes tense and person marking (3b), which are found on the lexical verb *ŋene-* 'to go' in the affirmative counterpart (3a). In negative sentences the lexical verb is obligatorily marked with a connegative (CON) form *-rA*, which in Uilta fuses with the verb stem (**ŋene-ri* > *ŋennɛɛ*, **ŋene-re* > *ŋennee*, Tsumagari, 2009, p. 10).

(3) Uilta (Tsumagari, 2009, p. 13)
   a. *ta.ri nari ŋennɛɛ-ni.*
      that man go.PRS-3SG
      'That man is going (somewhere).'
   b. *ta.ri nari* ***e-si-ni*** ***ŋennee.***
      that man NEG-PRS-3SG go.CON
      'That man is not going (anywhere).'

With some exceptions, languages from three of the four main branches of Tungusic preserve this pattern more or less intact (Hölzl, 2015). The major exception

among Tungusic languages is Jurchenic, where the Tungusic negative construction [*e-* v *-rA*][2] (e.g., Nedjalkov 1994 for details) has mostly been replaced with a new construction. Consider the following Manchu Example (4b) based on the negative existential *akū*, which fused with the imperfective participle *-rA*. This suffix *-rakū* replaces its affirmative equivalent *-mbi* (4a).

(4) Manchu (Written; constructed)
a. *te.re niyalma gene-mbi.*
that man go-IPFV
'That man is going (somewhere).'
b. *te.re niyalma gene-**rakū**.*
that man go-P.IPFV.NEG
'That man is not going (anywhere).'

This well-known development in Manchu has previously already been attributed to influence from Mongolian (e.g., Ikegami, 1999 [1979]), but to the best of my knowledge has never been studied from the point of view of CxG or including all attested varieties of Jurchenic. In fact, Manchu and Mongolian not only share this instance of constructionalization and an overall similar network of negative constructions, but also share similar constructional changes, which followed the emergence of the new negative construction.

Apart from Manchu, most Tungusic languages belong to Croft's type B. Above we have already encountered the Ewenic language Evenki as well as the Udegheic language Udihe. The Nanaic language Uilta has the negative verb *e-* seen above for standard negation (3b) and the negator *anaa* for negative existence and negative possession (5).

(5) Uilta (Tsumagari, 2009, p. 14)[3]
*tamaciga nari **anaa**.*
such person NEX
'There is no such man.'

(6) Manchu (constructed)
*tenteke niyalma **akū**.*
such person NEX
'There is no such man.'

2. Constructions will be shown with square brackets. As in this example, not all details will be included.

3. Optionally the negative verb may also be employed for negating the existential copula. Languages thus do not have to follow the negative existential cycle step by step.

Table 1 gives an overview of the negative existentials (NEX) and standard negation (NEG) in Jurchenic. Only the imperfective forms of the standard negator that are more frequent than the perfective forms (e.g., Written Manchu *-hakū*) are listed. The table also lists the Nanaic language Kilen (with contact to Manchu and Mandarin) as well as the two Ewenic languages Oroqen and Solon (with contact to Mongolic, Mandarin, and Manchu), which all have a similar albeit much more recent development. Oroqen and Solon otherwise preserve the original negative construction and Kilen has a verbal suffix that historically derives from the same source (Hölzl, 2015). These three languages will not be discussed any further here.

The negative existential may be reconstructed as **aq*^{h}*ʊa* for Manchuic. Only Jurchen^{B} and perhaps Yibuqi Manchu preserve an original final vowel, which was already lost in 17th century mainstream Manchuic. In Written Manchu, the negator is regularly written as *akū* (or *akû*) in the Möllendorff system. Recordings of modern dialects give the medial consonant as *q* ~ *k* and the final vowel as *u* ~ *o*. Verbiest (1682), who made some important obervations on the pronunciation of his time, recorded the form as *aco* (Latin <*c*> = [*k*^{h}]) in his *Elementa linguae tartaricae*. In order to clearly differentiate between constructs, "actual physical realizations of constructions", and constructions, "a generalization over constructs" (Fried, 2013, p. 423), the former will be indicated with the transcription used by individual authors (e.g., *akū* in von Möllendorff, 1892), but the latter will make use of an idealized form *aq*^{h}*ʊ*.

Yibuqi Manchu represents a special case as the negative existential *axuwo* ~ *axoꭥ* not only seems to preserve the final vowel and shows a spirantization of the plosive (7), on which also see Alchuka and Sibe below, but also allows the negator to directly attach to a verb stem (8) (cf. Written Manchu *mute-rakū* 'cannot').

Manchu (Yibuqi; Zhao Jie, 1989, p. 135)

(7) *ə.rə po ɳiʌɣmʌ* ***axuwo***.
this house person NEX
'There are no people (in) this house.'

(8) *pi mut'u-****axoꭥ***.
1SG able-NEG
'I cannot.'

But Yibuqi also has some examples of relics of a preceding participle (e.g., *-rxoꭥ*) and in general is very close to other Manchuic dialects.

Negation in Alchuka (*a'o*, *-ra'o*, *-xa'o*) is almost identical to Manchu (*akū*, *-rakū*, *-hakū*), but has lost the velar plosive that is present at least in reflexes in all of Manchuic.[4] The <'> probably indicates a hiatus as in *a'ɔ* = [*a.ɔ*] and <'> aspiration as in *-t'i* = [*t*^{h}*i*].

Table 1. Standard negators (imperfective only) and negative existentials in Tungusic. The transcription has been left unchanged and not all variants are shown

	NEX	NEX>NEG	NEG	Source
Jurchen[B]	**akua* 阿誇	**-rakua* 剌誇 **-rekua* 勒誇	–	Kane, 1989
Yibuqi Manchu	*axuwo* ~ *axoɷ*	*-axuwo* ~ *-axoɷ*	–	Zhao Jie, 1989
Written Manchu	*akū*	*-rakū*	–	von Möllendorff, 1892; Wuge & Cheng, 1730
Manchu (17th c.)	*aco*	*-raco*	–	Verbiest, 1682
Ningguta Manchu	*aku*	*-raku*	–	Aixinjueluo Yingsheng, 1990
Shenyang Manchu	*aku*	?	–	Aixinjueluo Yingsheng, 1991
Lalin/Jing Manchu	*aku*	*-r(a)ku*	–	Aixinjueluo Yingsheng, 1990
Sanjiazi Manchu	*ako*	*-lko*	–	Kim et al., 2008
Aihui Manchu	*aqo*	*-(r)qu* ~ *-(r)qo*	–	Wang Qingfeng, 2005
Sibe	*aqu*	*-(r)qu*	–	Chaoke, 2006
Alchuka	*aʾo*	*-raʾo*	–	Mu Yejun, 1986
Jurchen[A]	**asui* 阿隨	–	**ešin* 厄申	Kiyose, 1977
Bala	*asəi*	–	*-rʃən*, *ɛiwə*	Mu Yejun, 1987, 1988
Kilen	*atɕi*	*atɕi*	*-ɾtɕə*	Zhang Paiyu, 2013
Nanmu Oroqen	*aatʃin*	*-r(A) aatʃin*	*ə-* v*-r(A)*	Chaoke, 2007
Huihe Solon	*aaʃiŋ*	*-r(A) aaʃiŋ*	*ə-* v*-r(A)*	Chaoke, 2009

Alchuka (Mu Yejun, 1986, pp. 11ff.)

(9) *saŋə ta-tʻi* *puxɔ=i* *par* *katʃʻuxun* ***aʾɔ.***
bear beginning-ABL deer=gen towards harmonious NEG
'Bear and deer never got along well.'

4. But in one instance Mu Yejun (1986, p. 16) mentions a form that preserves a velar fricative and lacks the initial consonant (ɔ.tʃi-raxɷ).

(10) *kə.rut'uri=mə ufara-t'i kɔ.tʃi-**ra'ɔ**.*
this chance=ACC lose-COND become-P.IPFV.NEG
'This opportunity must not be missed.'

Examples (9) and (10) illustrate interesting differences to Manchu, such as the lack of palatalization (Manchu - ***ci*** 'ABL; COND', ***u**c**u**ri* 'opportunity'), different lexical items (Manchu *lefu* 'bear'), a variable accusative case marker with the variants *=p(ə)~ =m(ə)~ =w(ə)~ =u* (at least Written Manchu invariably has *=be*), as well as a problematic initial consonant *k-* (Manchu *acuhūn* 'harmonious', *e.re* 'this', *o-* 'to become', e.g. Hölzl 2017). These are some reasons for differentiating Alchuka from Manchuic (see Mu Yejun 1986, 1987). But, apart from phonological differences, the expression *acuhūn akū* 'unharmonious' as well as the construction [V-*ci o.jo-raq*^{h}*ʊ*] 'must not' (a combination of standard negation with [V-*ci o-*] 'be possible') are identical in Written Manchu, including the irregular stem extension of *o-* 'to become, to be'.

The negative existential in Balaic must go back to **asui* but may ultimately be connected with the other Tungusic negative existentials, all of which start with *a~*. Balaic also differs from the other two branches in that the development of the new negation strategy did not take place. Jurchen^{A} and possibly Bala as well preserve traces of the Tungusic negative verb, but there is no complete match as there is in the negative existential.[5] Thus, by definition, the development of the new negative construction cannot yet have happened in Proto-Jurchenic. It may have started in Manchuic and spread to Alchukaic, though this remains speculative. Bala apparently has adopted the Nanaic negative strategy through Kilen instead, e.g. Manchu *gene-rakū* 'don't/doesn't go', but Bala *ənə-rʃən* (Mu Yejun, 1988, p. 11). The suffix *-rʃən* (and variants) in Kilen goes back to a construction in which the negative verb followed the lexical verb, i.e. *ənə-* 'to go' + *-r(A)* 'CON' + (**ə*)*-ʃə-n* '(NEG)-PRS-3SG' (Hölzl, 2015). This is a direct cognate with Uilta *e-si-ni ŋennee* (< **ŋene-re*) in (3b), in which the negative verb precedes the lexical verb. In this example the entire expression, including the verb *ənə-*, has been adopted from Kilen (cf. Manchu ***g**ene-*, Uilta ***ŋ**ene-*).[6] Balaic will not be further addressed here.

The functional side of the negative existential is not easy to grasp. But based on data from the two standard dictionaries of Manchu (Hauer, 2007; Norman, 2013) one may assume a network of related constructions (also see Gorelova 2002). As we will see in Section 3, this network has close parallels in Mongolic

5. Jurchen*A* also has a negator **eihe* 厄一黑 'not yet' that might be related to Bala *ɛiwə*.

6. Jurchen*A* *ešin is cognate with Uilta *e-si-ni* as well, but the form of the negator in combination with the verb in Bala makes a borrowing from Kilen much more likely than an autochthonous development.

(e.g., Janhunen 2003a, p. 98). The most central node for this network is the schema [N $aq^hʊ$], e.g. *morin akū* 'without a horse, horseless'. This is the basis of formulaic expressions such as *baitakū* 'no problem' (cf. *baita* 'matter') or fixed collocations such as *derakū* 'shameless, undignified' (cf. *dere* 'face'). These are the result of "lexical" constructionalization and subsequent fusion. Combined with the fully schematic attributive construction [ATTR N], e.g. *sain acabun* 'good omen', there is the extended schema [[N $aq^hʊ$] ATTR N], e.g. *gebu akū bithe* 'an anonymous letter, a letter without name'. An example from spoken Sanjiazi Manchu illustrates that a case marker may attach to the noun, but marks the whole noun phrase (11).

Manchu (Sanjiazi; Kim et al., 2008, p. 166)
(11) [[*molin* ***ako***]*ǰelin*]=*de*
horse NEX reason=DAT
'because of not having a horse'

This extended construction in turn is the origin of fixed lexicalized expressions such as *gebu akū simhun* 'ring finger', which literally means 'finger without name'. A case marker may also directly attach to *akū*, but marks the whole subordinate clause.

Manchu (Written; Hauer, 2007, p. 129, from the Jinpingmei)
(12) [*si.n-i beye elhe* ***akū***]=*be bi sa-****rkū****.*
2SG.OBL-GEN body healthy NEG=ACC 1SG know-P.IPFV.NEG
'I did not know that [you don't feel well].'

Notice, too, the reduced form *-rkū*, which can be found exclusively with the verb *sa-* 'to know'. Host class expansion (e.g., Trousdale, 2014), here understood as the increase in schematicity of one part of the construction, is responsible for the fact that *akū* in Example (12) attaches to an adjective rather than a noun. The construction [N $aq^hʊ$] thus becomes more schematic and allows the pattern [ADJ $aq^hʊ$], e.g. *sain akū* 'to be not good'. In this construction *akū* has the function of a negative copula. Again, there are some lexicalized and contracted forms such as *elhekū* (< *elhe akū*) 'not (feeling) well'.

In Manchu, existence and possession are expressed with an intransitive locational construction [N_X *-de* N_Y *bi(-)*], which literally means 'at X is/are Y' (> 'X has Y'). The dative case marker *-de* in Manchu also has the function of a locative. If integrated with this clausal existential or possessive construction, the negative existential replaces the copula or existential verb *bi(-)*, i.e. [N_X *-de* N_Y $aq^hʊ$]. Consider the following question-answer pair from Sibe, where the dative has the form *-d* (13a, 13b).

Sibe (Chaoke, 2006, p. 222)

(13) a. *ʂi.ni-**d*** *untʂu* *gəv* ***bi**-m=na?*
2SG.OBL-DAT different name EX-IPFV=Q
'Do you have a different name?'

b. *mi.ni-**d*** *untʂu* *gəv* ***aqu**.*
1SG.OBL-DAT different name NEX
'I don't have a different name.'

One may also have answered the question with Example (14).

Sibe (Chaoke, 2006, p. 226)

(14) ***aqu** ,mi.ni-d* *aqu.*
NEG 1SG.OBL-DAT NEX
'No, I don't have (one).'

This illustrates that [$aq^hʊ$] may also function as a negative one-word answer meaning 'no'. This use without a preceding noun may be seen as the result of an ellipsis that often affects answers. The context of question-answer pairs is also the origin of yet another expansion of the network. The form *akūn* in (15) is the result of fusion with the question marker *=ni* in negative alternative questions.

Manchu (Written; Hauer, 2007, p. 10, from the Jinpingmei)

(15) *aika baita* *bi-he=o*, ***akūn?***
any matter EX-PFV=Q NEX.Q
'Has there been any matter or not?'

Ellipsis also explains why *akūn* is not preceded by a noun in this case. In alternative questions most languages allow the ellipsis of identical parts, in this case *aika baita*. Given that to my knowledge the question marker *=ni* has been lost in most modern dialects, the form exists exclusively in Written Manchu and a few Manchurian dialects. Sibe, for example, generally has the still analyzable sentence final particle *=na*, e.g. *gia-q^{ú}=na?* 'take-NEG=Q' (Zikmundová, 2013, p. 82).

The development of standard negation is nothing but an extension of the basic schema [N $aq^hʊ$] to [V-PTCP $aq^hʊ$], possibly influenced by [ADJ $aq^hʊ$]. Participles are similar to nouns in that they may take case markers but also similar to adjectives in that they may function as relative clauses (attributes). Perhaps the schematization to allow the negation of adjectives made the development to standard negation more natural. Manchu has two main participle forms, imperfective *-rA* and perfective *-hA*, both of which can be negated with *akū*. The development in Manchu also encompasses an instance of so-called verbalization, the reanalysis of a nominal as a verbal predicate (Malchukov, 2013). The following is an example from the Ewenic language Even.

Even (Malchukov, 2013, pp. 181f.)

(16) a. *bej* [*bujusemŋe*] [*bi-si-n*].
man hunter COP-PST-3SG
'This man was a hunter.'

b. *bej* [*hör-če*] [*bi-si-n*].
man go-P.PFV COP-PST-3SG
'This man had left.'

c. *bej* [*hör-če bi-si-n*].
man go-P.PFV AUX-PST-3SG
'This man had left.'

Similar to Manchu, the first step is schematization or host class expansion. In this case, the copula *bi-* that marks an equational clause in (16a), may now also follow a participle form in (16b) (i.e., a nominalized verb). The next step is verbalization in the strict sense, with reanalysis of the participle as a verbal instead of a nominal form. The copula now fulfills the function of an auxiliary (16c). Subsequently, the newly formed complex verb form may merge. Table 2 gives an overview of the gradual diachronic changes observed in (mainstream) Manchuic following the host class expansion and the reanalysis. These changes may be described as bonding and phonetic erosion (i.e. constructional changes).

Table 2. Idealized diachronic development of the standard negator in (mainstream) Manchuic

Time	NEG	Description
?	**-rA* *aq^{h}ʊa*	unattested, Manchu *-rA* has vowel harmony *-ra* ~ *-re* ~ *-ro*
~1500	**-rAq^{h}ʊa*	Jurchen^{B} (Kane, 1989), reduced vowel harmony **-ra* ~ **-re*, e.g. **sa-rakua* 撒剌誇 'don't know'
~1600–	*-raq^{h}ʊ*	Written Manchu *-rakū*, e.g. *mute-rakū* 'cannot', exception *sa-rkū* 'don't know', almost no vowel harmony
~1700	*-raq^{h}ʊ*	spoken and written Manchu given as <*-raco*> by Verbiest (1682), exception <*sa-rco*> 'don't know'
~1900	*-r(a)q^{h}ʊ*	the variant [v-*rq^{h}ʊ*] can be found in written (Stary, 2009, p. vii) and spoken (Kałużyński, 1977, passim) Manchu in Sibe context, i.e. extension of exceptions, e.g. *mute-rkū* 'cannot'
~2000	*-(r)q^{h}ʊ*	e.g. Sibe *sa-qu*, 'don't know', *mutu-qu* 'cannot' most dialects preserve reflexes of the *r*

Interestingly, synchronic variation until today shows some indications of the earliest stage **-rA aq^{h}ʊa*, which is otherwise unattested (17, *-rA* > *-l*).

Manchu (Sanjiazi; Kim et al., 2008, p. 166)
(17) *əniŋ jo-m mutu-**l** **ako** o-xo.*
today go-CVB.IPFV can-P.IPFV NEG be-P.PFV
'I could not go today.'

The innovative last step, during which the rest of the participle marker was lost, can be seen in Sibe (18) and sometimes in the Aihui dialect as well (19).

Sibe (Chaoke, 2006, p. 221)
(18) *bi dzəŋkən sa-**qu**.*
1SG really know-NEG
'I really don't know.'

Manchu (Aihui; Wang Qingfeng, 2005, p. 241)
(19) *bu! bi ɛli-n ɢia-**qu**.*
NEG 1SG endure-CVB.IPFV take-NEG
'No! I don't accept/can't stand it.'

Clearly it is the imperfective participle *-rA* that was first introduced into the newly formed schema [V-PTCP $aq^hʊ$]. Note that *-rA* is also the form found as connegative in the Tungusic negative construction [*e-* V-*rA*] as well as in the Manchu prohibitive construction [*ume* V-*rA*]. Over time, the construction became more schematic and allowed the use of the perfective participle *-hA* as well. Thus, the combination with the perfective participle must be a secondary innovation, and in fact it is less frequent even today. In addition, it preserves some restricted vowel harmony that was almost completely lost in *-rakū* and shows fewer signs of erosion. Consider an example from Aihui Manchu (20).

Manchu (Aihui; Wang Qingfeng, 2005, p. 60)
(20) *i bo-də dojə-**χa-qu**.*
3SG house-DAT enter-P.PFV-NEG
'(S)he did not enter the house.'

Given that Aihui Manchu allows the negative suffix *-qu* ~ *-qo* (= Manchu *-rakū*) to attach to plain verb stems, the sequence *-χa-qu* (= Manchu *-hAkū*) has here been analyzed as two suffixes. More frequently than with the participle *-rA*, there are examples in which *-hA* and the negative existential have not fused into one form (e.g., 21).

Manchu (Yibuqi; Zhao Jie 1989: 126)[7]
(21) *kuᴀ niᴀɣmᴀ xɛ tɕi-**ɣə** **axoɷ**.*
other person still come-P.PFV NEG
'The other person has not come yet.'

7. This example is additionally based on the Chinese construction *[hái méi* V *guò]* 还没V过 'not yet, never' of which *hái* 'still' has been borrowed as *xɛ*.

A count in a short war diary produced in Written Manchu towards the end of the 17th century reveals the numbers listed in Table 3. The form *-rakū* is overwhelmingly more frequent than all other variants and shows less synchronic variation, which is a sign that it is older and more strongly conventionalized than the forms based on the perfective participle *-hA*. The perfective form *-hAkū* furthermore has the irregular forms *-kAkū* and *-ngkAkū* for some verbs, which were not attested in this small scale corpus.

Table 3. Frequencies of negative forms in a 17th century war diary (9160 tokens) (Di Cosmo, 2006)

Form	Frequency
-rakū(-)	37
-ha akū	7
-hakū(-)	5
-hekū	3
-he akū	1
-ka akū	1
-rkū	1

Interestingly, the *Qingwen qimeng*, the first Manchu grammar in Chinese (Wuge & Cheng, 1730, p. 27b), states that the form *-hakū* and its variants necessarily have the vowel *a*. Perhaps this is intended to be prescriptive rather than descriptive but may also indicate a development towards stronger bonding and less analyzability. In fact, the form *-hakū* ~ *-ha akū* is more frequent than *-hekū* ~ *-he akū*. The vowel harmonic form *-ho akū* is rare and *-hokū* seems to be almost entirely absent.

Manchu has no other participle marker that may have been negated by the negative existential. Alchuka has a participle *-ʑï* (< **-si*), however, which seems to have imperfective meaning similar to Manchu *-rA*.[8] It may take case markers and appears in the Alchuka prohibitive construction [*ɔmə* v-*ʑï*], which corresponds to [*ume* v-*rA*] in Manchu and [*edʒi* v-*rA*] (singular) in most other Tungusic languages (Mu Yejun, 1986; Hölzl, 2015).[9] There is no apparent reason why this par-

8. Cf. the irregular stem extension in *bi.si-re* 'COP-P.IPFV' in Written Manchu, the form *e-si-ni* in Example (3b) from Uilta, and the form *bi-si-n* in Example (16) from Even.

9. Note that in some other Tungusic languages *-si* is the connegative form of the copula *bi-* (Hölzl, 2015, p. 146).

ticiple may not function as host for the negative existential, unless the standard negative construction has been adopted from Manchuic.

Table 4 shows the most important Manchu verb forms. There is an asymmetry, because the imperfective (dictionary form) *-mbi* cannot be negated. Finite verb forms in Manchu usually contain either the existential (copula) *bi(-)* or the negative existential *akū* (cf. Sinor, 1949). An exception is *-hA*, which may sometimes act as a finite form. Perhaps the negation of *-mbi* is excluded because it is too strongly fused to allow a clear analysis (there is no suffix **-m*) but still transparent enough to allow an identification of the copula *-bi*.[10] The existential, of course, cannot be negated because it has the suppletive negative form *akū*. This may have been one reason why the participle was negated instead.

Table 4. Asymmetry under negation in Written Manchu (cf. Gorelova, 2002, passim)

Meaning	Assertion	Negation	Changed function
IPFV	*-mbi(-)*	–	–
PST	*-hA-bi(-)*	*-hAkū-bi(-)*	–
P.IPFV	*-rA*	*-rakū*	+finite
P.PFV	*-hA*	*-hAkū*	(+finite)

In order to express temporal or aspectual categories, both the standard negation construction and the negative existential construction allow an integration with the auxiliary construction [V AUX-] (22, also see Example 17).

Manchu (Written; Hauer, 2007, p. 51, from the Jinpingmei)

(22) *mi.n-i* *ahūn* *e.re gese nime.ku* ***akū bi****-he.*
1SG.OBL-GEN e.brother this like illness NEG be-PFV
'My elder brother did not have such an illness.'

An extension of this combination of negative existential and auxiliary is the fixed expression *akū o-* 'to die' (often used with perfective aspect *-ho*), which literally means 'to become non-existent'. A development that is restricted to Sibe led to the emergence of several suffixes that may attach to the negative suffix. In a recording made in September 2014 in Qapqal, Written Manchu *mute-rakū o-ho* 'can-P.IPFV.NEG become-P.PFV' was read as one word, [mʊtʰʊ̆ˈχʊːə] or [mʊtʰʊ̆ˈχʊːʌ], by a native Sibe speaker. But there are also forms that are more clearly analyzable (23).

10. In other Jurchenic varieties the connection to *bi* has been obscured by phonological changes, e.g. Sibe *-m,* Alchuka *-mei,* Bala *-mi.*

Sibe (Zikmundová, 2013, p. 132)

(23) *śi.m-b* *sa.v-ʁa-**qů**-maq* *yav-ʁə.*
2SG.OBL-ACC see-P.PFV-NEG-CVB.PFV go-P.PFV
'I did not see you there and so I left.'

The converb *-maq* is specific to Sibe. Other examples mentioned by Zikmundová (2013, pp. 132f.) are *ji-**qů**-mə* 'come-NEG-CVB.IPFV' (Manchu *ji.de-rakū o-me*) and *jə-**qů**-či* 'eat-NEG-COND' (Manchu *je.te-rakū o-ci*). These examples illustrate the emergence of suffixes out of the inflected auxiliaries as well as the reduction of irregular stem extensions.

Standard negation, too, has led to some separate sub-nodes within the network. For instance, there is a comparative construction [N_X N_Y *-de isi-raq^{h}ʊ*], which literally means 'X does not reach Y' (24).

Manchu (Gorelova, 2002, p. 157)

(24) *i* *mi.n-de* *isi-**rakū**.*
3SG 1SG.OBL-DAT reach-P.IPFV.NEG
'I am better than him/her.'

Figure 3 summarizes the previous discussion and schematically shows the network of related constructions in Manchu. Apart from the newly formed standard negation construction, there are several instances of new micro-constructions (atomic and complex) that are the result of "lexical" constructionalization (Traugott & Trousdale, 2013, p. 192).

Several properties of standard negation, such as the fact that it can take case markers, may function as an attributive form (relative clause), and has the interrogative forms *-rakūn* and *-hakūn*, were inherited from the source schema and are thus not listed separately.

The scope of negators over different categories of negation may be shown with the help of a conceptual space (aka semantic map) (Veselinova, 2015; Hölzl, 2015). For the purposes of our discussion, the two categories of standard negation (NEG) and negative existence (NEX) are sufficient (Figure 4). Since the original standard negation construction in Manchuic is unknown, the Proto-Tungusic construction is shown instead.

This suggests that Manchu did not actually follow Croft's cycle until the end. Instead, the last step may indicate that some varieties such as Sibe today have reached type B again, with a negative existential *aqu* that is formally different from the standard negator *-qu*. But at least the Aihui dialect has reached type C ~A and allows the negation of the existential *bi-* with the newly formed negator, i.e. *bi.dzə-qo* (Wang Qingfeng, 2005, p. 114, standardized). Note the same irregular

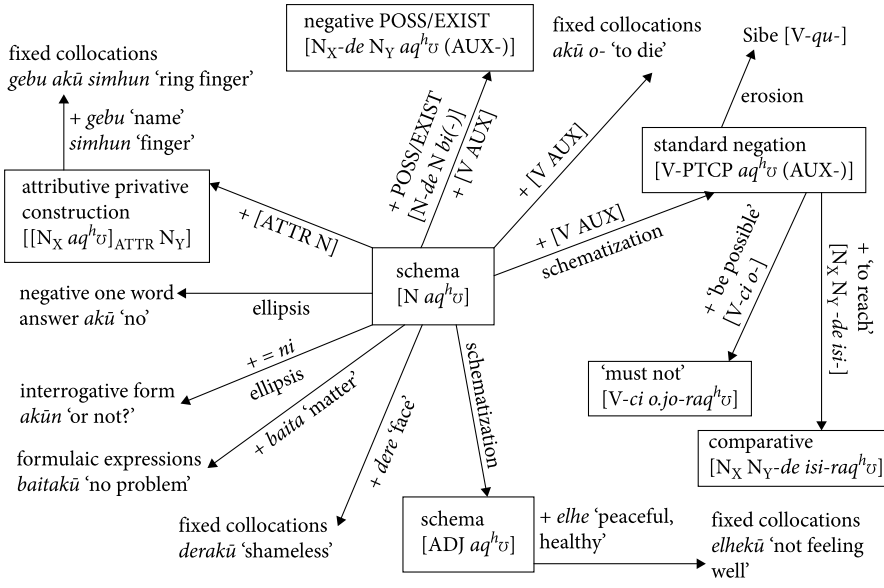


Figure 3. Partial network of the negative existential construction in Manchu mostly based on data in Hauer (2007) and Norman (2013). A plus sign means 'integration with'. Arrows illustrate the extension of the network. Boxes indicate productive schemas

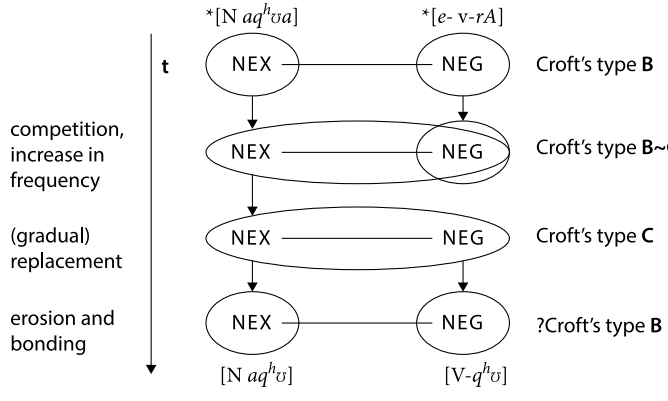


Figure 4. Changes in the semantic scope of negators in Manchuic

stem extension but the absence of the participle as compared to Written Manchu *bi.si-re* (= Aihui *bi.dzə-rə*).

Finally, let me briefly address the etymology of the Manchuic existential negator **aq^{h}ʊa*. A recent attempt by Robbeets (2014, p. 406) to explain Tungusic negative existentials in general runs into several problems. First of all, the alleged

Proto-Tungusic reconstruction was sometimes given as **ana-* and sometimes as **a:na-* with a long vowel. Without an explanation one can only assume that one of the forms is a typographical mistake or that the vowel length is assumed to be variable. Furthermore, the derivations in individual languages are highly implausible, e.g. Manchu *aku*: < **a(:)na-xU*. Note that the Manchu form is cited in an incorrect way, i.e. Manchu *aku*: instead of *akū*, where the diacritic marks vowel quality and not quantity. Several forms are missing altogether (e.g., Nanai *aba*, Bala *asəi*), and one form, Udihe *a-ta-*, is not an existential at all but represents a future form of the negative verb with an additional vowel change (cf. Evenki *e-te-*, Negidal *a-ta-*, Hölzl, 2015, p. 123). The negative existentials, all of which start with an *a~*, might be related to each other, but one cannot exclude the possibility of relabeling or chance resemblance. Interestingly, Manchu has the verbs *akū-nji-* 'to arrive at this shore, to arrive at this side', and *akū-na-* 'to reach (the opposite shore), to go to the end'. These two contain the directional suffixes *-nji* 'to come to do' and *-na* 'to go to do'. The verbs express motion and have a semantic and collocational connection to rivers, e.g. *dalin-de akū-na-* 'to reach the shore' (*dalin-de* 'shore-DAT'). Note that motion is a frequent source domain for death and absence, which in turn may be the origin for negative existence. The original meaning of *akū(-)* may thus have been 'towards the (other) shore'. But to verify this tentative etymology more in-depth research in other Tungusic languages is necessary.[11]

3. Mongolic

According to one classification (Janhunen, 2006, p. 232, 2012a, pp. 2ff.), Mongolic has four main subgroups, shown in (25). The recently discovered language Kangjia has been added. Most Mongolic languages are mainly spoken in Mongolia and in the adjacent regions in China and Russia. Only Kalmyk, an aberrant dialect of Oirat in Europe, and Moghol in Afghanistan are located outside of this area. All Mongolic languages combined have about 5 to 7 Mio speakers (Janhunen, 2006, p. 233). Only Moghol is perhaps extinct.

(25) 1. Dagur(ic)
1.1 Dagur

11. The verb *akū-* 'to die < to be(come) non-existent', together with its conditional form *akū-ci* 'if not, otherwise', its causative/passive form *akū-mbu-* 'to use up, to exhaust, to endeavor', etc., is most likely a secondary derivation from the negative existential *akū*.

2. Central/Common Mongolic
   2.1. Khamnigan Mongol
   2.2. Buryat
   2.3. Khalkha Mongolian
   2.4. Khorchin Mongolian
   2.4. Ordos
   2.5. Oirat
3. Shirongolic
   3.1. Shira Yughur
   3.2. Huzhu Mongghul
   3.3. Minhe Mangghuer
   3.4. Bonan
   3.5. Kangjia
   3.6. Santa
4. Moghol(ic)
   4.1. Moghol

While Tungusic probably has a time-depth of about 2000 years, Mongolic is only about 800 years old (e.g., Janhunen, 2003a, 2012b). However, Mongolic is part of an older and larger language family, for which the name Khitano-Mongolic has been proposed (Janhunen, 2012c, p. 114). Not much is known about negation in Khitan, however, which is the main representative of the extinct other branch (also referred to as Para-Mongolic) of the Khitano-Mongolic language family. Here the focus will lie on parts of Common Mongolic.

Proto-Mongolic had a distinction between the two preverbal negative particles **ese* for past and perfective forms and **ülü* for non-past and imperfective forms (Janhunen, 2003a, p. 27; Brosig, 2015, p. 71). These negators are retained in some languages, but have been replaced with the negative existential **üge(y)i* in several others (Table 5). In some languages there is competition between the old and the new forms. Languages without the development will be excluded here as will combinations of the negative existential with converb forms. Apart from Common Mongolic, Dagur and the two Shirongolic languages Mangghuer and Bonan also have similar changes. The latter two will not be addressed any further (see Brosig, 2015). According to Róna-Tas (2017, p. 145), the Khitan negator **gi* might be related with Mongolic **üge(y)i* as well, although many details have yet to be explained.

In general, the development in Mongolic is rather recent, which is why it is historically attested as well. Brosig (2015, p. 73), based on data by Toγtambayar, has shown how the relative frequency of tokens has shifted from *ese* and *ülü* to *ügei* over the course of the last three and a half centuries. The aberrant Oirat dialect Kalmyk was separated from the Mongolic speaking areas in the early

Table 5. Standard negators and negative existentials in Mongolic (cf. Brosig, 2015, p. 70). The transcription has been left unchanged and not all variants are shown

Language	NEX	NEX>NEG	NEG	Source
Khalkha	*ügüi*	*=güi*	–	elicited
Khorchin	*uguɛ*	*=guɛ*	–	Brosig, 2015
Ordos	*ügüei*	*=ügüei*	–	Georg, 2003
Oirat	*ugä, …*	*=goo*	*ül(ü), es(e)*	Birtalan, 2003
Buryat	*ügii*	*=güi*	–	Skribnik, 2003
Khamnigan	*ugui*	*=gui*	–	Janhunen, 2003b
Dagur	*uwei*	*=wei*	*ul, (es)*	Tsumagari, 2003
Mangghuer	*gu-*	*gu-*	*(l)ai, sai*	Brosig, 2015
Bonan	*ʻgi-*	*ʻgi-*	*elə, esə*	Brosig, 2015

17th century, but nevertheless has the same pattern. Consequently, the development must have started earlier than that, but the details have yet to be investigated (Brosig, 2015, p. 74). Among Jurchenic languages, the oldest attestations of the development are found in Jurchen[B] that were perhaps recorded around 1500 (Kane, 1989, p. 99f.), when the new construction had already completely replaced the old one. This might indicate that the change happened earlier in Jurchenic than in Mongolic, but in general Mongolic exerted a much stronger influence on Tungusic than vice versa.

The negated forms in Mongolic are more analyzable than in Jurchenic, but in some modern Manchuic dialects at least the perfective forms are synchronically analyzable as well, and there are contracted forms in Mongolic, too (Brosig, 2015, p. 75). In the following example from Dagur (26), *-gw uwei-* may also have the less analyzable form *-wwei-*, which is analogous to Manchu *-rakū*.

Dagur (Tsumagari, 2003, p. 150)[12]

(26) *naadjaa-gw **uwei**-sul.*
play-P.FUT NEG-3PL
'They are not playing.'

As in Manchu, the fact that only participles could be negated by the newly created construction led to asymmetrical patterns of the following type (27a, 27b).

12. This example illustrates person marking on the negative existential that is absent in most of Mongolic and Jurchenic.

Mongolian (Janhunen, 2012a, p. 252)

(27) a. *te.r oc-**e.n**'*.
3SG go-DUR
'(S)he will go.'

b. *te.r oc-**ex=gwai**.*
3SG go-P.FUT=NEG
'(S)he will not go.'

While the asymmetry is analogous, Mongolian has a larger inventory of forms than Manchu (Table 6). Perhaps due to the need to express the meaning of finite verb forms under negation, the negated participle forms have acquired different meanings than their affirmative counterparts.

Table 6. Asymmetry under negation in Mongolian (Janhunen, 2012a, pp. 157, 180, 221, 252)

Meaning	Assertion	Negation	Changed Meaning
PRS.IPFV (DUR)	*-n- ~ -n=AA ~ -e.n'*	–	–
PRS.PFV (CONF)	*-l- ~ lAA*	–	–
PST.IPFV (TERM)	*-b- ~ -b=AA ~ -e.b ~ -eb*	–	–
PST.PFV (RES)	*-J- ~ -Jai ~ -e.j*	–	–
P.FUT	*-ex*	*-ex=gwai*	PRS
P.IPFV	*-g.AA*	*-g.AA=gwai*	PRT.IMPRF
P.PFV	*-sen*	*-seng=gwai*	PRT.PRF
P.HAB	*-deg*	*-deg=gwai*	–

In Buryat the asymmetry has been resolved by an expansion so that not only all participles but also all finite verb forms can be negated with the new negator.

Buryat (Shineken; Yamakoshi, 2007, p. 13)

(28) a. ***muŋgu=gui***
silver=NEX
'without silver'

b. *oj-**aa=gui*** *nɔxɔi*
tether-P.IPFV=NEG dog
'dog(s) that are not tethered'

c. *jab-**na=gui**=b^{j}.*
go-DUR=NEG=1SG
'I do not go.'

These three Examples (28a, 28b, 28c) illustrate a step-by-step increase in schematicity. Note that the negative existential may also have the abbreviated form

=*gui* following nouns, which is unlike Manchu (cf. *menggun akū* 'without silver'). For a discussion of its morphosyntactic status as either suffix or enclitic see Yamakoshi (2007, p. 14), Janhunen (2012a, p. 109), and Brosig (2015, p. 75).

The network of constructions in Mongolian is almost completely analogous to that in Manchu (see Figure 2 above, Janhunen 2003a: 98). Negation in Mongolic has often been investigated (e.g., Brosig 2015 and references therein), but in order to make the comparison with Jurchenic as clear as possible, the following data from Cyrillic Khalkha were elicited in April 2016 from a native speaker living in Germany. The analysis follows Janhunen (2012a). The negative existential has the form *ügüi*/ҮгҮй, which may function as a negative one-word answer 'no' if used alone. As in Buryat there is the reduced form =*güi*/гҮй that enclitically attaches to nouns as well (29).

(29) морьгҮй
mory=***güi*** horse=NEX
'without a horse'

(30) нҮҮргҮй
nüür=***güi*** face=NEX
'shameless'

(31) юмгҮй
yum=***güi*** thing=NEX
'nothing'

Example (30) is a fixed collocation corresponding to Manchu *derakū*. The last Example (31) is a fixed expression similar to, but with a different meaning than, Manchu *baitakū* < *baita akū* 'no problem'. However, Dagur has borrowed Manchu *baita* 'matter' and has a comparable expression, e.g. Tacheng Dagur *bajt uwee* (Yu et al., 2008, p. 177). As in Manchu there is an attributive construction [[N =*güi*] $_{\text{ATTR}}$ N] (32), and some fixed collocations similar to Manchu *gebu akū simhun* 'ring finger' (33).

(32) хҮҮхэдгҮй хҮн
xüüx-e.d=***güi*** *xün*
child-PL=NEX man
'a man without children'

(33) нэргҮй хуруу[13]
ner=***güi*** *xuruu*
name=NEX finger
'ring finger'

In negative alternative questions it may take the question marker *=(y)UU* (34). Similar to Sibe, but unlike Written Manchu, the negator and the question marker did not fuse.

(34) чи өчигдөр намайг харсан уу, үгүй юу?
*ci öcigdör nam-ai.g xar-san=uu, **ügüi**=y.uu?*
2SG yesterday 1SG.OBL-ACC see-P.PRF=Q NEG=Q
'Did you see me yesterday or not?'

There is also a construction meaning 'to die' with a form analogous to Manchu *akū o-* 'to become non-existent' (35).

(35) тэр үгүй болсон.
*te.r **ügüi** bol-son.*
3SG NEX become-P.PRF
'(S)he died.'

A slight difference to Manchu is the fact that adjectives may be negated with the negative copula (36) as well as the negative existential (37). The difference appears to be a focus on the clause in the former and on the adjective in the latter (cf. Yamakoshi, 2011, p. 169). There are collocations similar to Manchu *elhekū* 'not (feeling) well' (38).

(36) сайн биш
*sain **bish***
good NEG
'not good'

(37) сайн үгүй
*sain **ügüi***
good NEG
'not good'

(38) энхгүй
*enx=**güi*** peaceful=NEG
'not peaceful'

However, there are at least some instances in which an adjective can enter the Manchu negative copula construction [N *waka*]. Another difference to Written Manchu is the possibility for *=güi* to attach not only to participle forms of verbs

13. This expression is not very common but has been included here for its similarity with Manchu.

(39) but to the existential copula (40) as well (Croft's type A). However, as seen before, this is also a possibility in Aihui Manchu (type C~A).

(39) би яваагүй байсан.
*bi yaw-aa=**güi** (bai-san).*
1SG go-P.IMPRF=NEG COP-P.PRF
'I did not go.'

(40) надад ном байхгүй.
*nad-a.d nom **bai**-x=**güi**.*
1SG.OBL-DAT book COP-P.FUT=NEG
'I don't have a book.'

The similarity between Manchu and Mongolian goes so far that even extremely specific derivations of standard negation for negative obligation (41) and comparison (42) have analogous forms.

(41) чи ирвэл болохгүй.
*ci ir-**wel** **bol**-o.x=**güi**.*
2SG come-COND become-P.FUT=NEG
'You must not come.'

(42) тэр надад хүрэхгүй.
*te.r nad-**a.d** **xür**-e.x=**güi**.*
3SG 1SG.OBL-DAT reach-P.FUT=NEG
'I am better than him/her.'

There is a complete correspondence to the constructions in Manchu. The major difference concerns the phonological shape of the conditional *-ci*, the dative *-de*, and the verbs *o-* 'to become' as well as *isi-* 'to reach.

Despite frequent claims to the contrary, a genetic relationship between Tungusic and Mongolic has not been proven. There is still the possibility that such a connection may be found in future investigations, but for the moment they must be treated as separate families that influenced each other. In most cases discussed here the direction of influence certainly went from Mongolian to Manchu. In fact, in several examples given above, some words (or at least their phonetic shape) are derived from Mongolic, too (e.g., Manchu *morin* 'horse', *sain* 'good', *elhe* 'peaceful', *menggun* 'silver'). There are many additional similarities not mentioned yet and several surrounding languages share parts of this network of constructions as well. For instance, Manchu *baitakū* 'no problem' is most likely a *calque* of Mandarin *méi shì* 没事 'NEX matter'. But a detailed investigation goes beyond the possibilities of

this paper. For reasons of space the following section will give a bird's eye view of similar developments from negative existentials to standard negators, exclusively.

4. Surrounding languages

Chinese Pidgin Russian has two negators, preverbal *ni* (Russian *не*/*не*) and the more frequent postverbal existential *netu* (Russian *нет(у)*/*нем(у)*), the syntactic behavior of which may have been influenced by Tungusic or Mongolic. The negator *netu* is not only a negative existential (43a), but can also function as the standard negator (43b).

Chinese Pidgin Russian (Shapiro, 2010, p. 30)

(43) a. *synka* ***netu***.
son NEX
'(I) don't have a son.'

b. *əta liba pomilaj* ***netu***.
this fish die NEG
'This fish did not die.'

Dunan, like most Japonic languages, by default uses a negative suffix *-anu-*, but the negative existential *minu-* (44a) combines with a medial verb form to express the negative perfect (44b).

Dunan (Yamada, Pellard & Shimoji, 2015, pp. 469, 471)

(44) a. *khumi=ja baga nnani=ja* ***minu****-n.*
here.LOC=TOP 1.GEN clothes=TOP NEX-IND
'My clothes are not here.'

b. *madi Tharu=nki thuras-****i*** ***minu****-n.*
yet PN=DIR give-MED NEG-IND
'(I) have not given (it) to Taro yet.'

The Amuric (Janhunen, 1996) language Nivkh has a negative existential *q'au-* (45a) that may also function as the standard negator (45b). In the latter case the verb has to take the specific nominalizing suffix *-roχ ~ -toχ ~ -doχ*. The following examples are from the dialect spoken along the lower Amur.

Nivkh (Amur; Nedjalkov & Otaina, 2013, pp. 96f.)

(45) a. *ñ-uin hə=bityə* ***q'au****-ḍ.*
1SG-LOC that=book NEX-IND
'I don't have this book.'

b. *if p'rə-**doχ q'au**-ḍ.*
3SG come-N NEG-IND
'(S)he did not come.'

Apart from this construction, Nivkh has several more ways of expressing negation.

The Uralic language Selkup, spoken in Siberia to the northwest of Mongolia, uses the preverbal negator *ašša*, allegedly of Tungusic origin, for the negation of the present tense, but a postposed negative existential (46a) in the past tense (46b).

Selkup (Northern, Taz; Wagner-Nagy, 2015, pp. 144, 137, from Kuznecova et al.)

(46) a. *ukkɨr poo ämtä **čääŋka**.*
one tree at.all NEX.3SG
'There aren't any trees at all.'
b. *man ilɨ-**ptɔɔ**-mɨ **čääŋka**.*
1SG live-N-1SG NEG.3SG
'I did not live.'

Note the nominalization in the second example, which literally means 'My living does not exist.'

Still more similar to Jurchenic and Mongolic are patterns observed in some Turkic languages of the region. Kazakh (Muhamedowa, 2016), for example, has a negative verbal suffix *-MA* as in *kel-**me**-gen-min* 'come-NEG-PERF-1SG'. But there is also the possibility of employing the negative existential *žoq*, e.g. *kel-**gen** **žoq**-pïn* 'come-PERF NEG-1SG'. The negative existential follows a participle form (*-GAn*), while the negative suffix attaches directly to the verb. In Kazakh, the semantic difference between *-MA* and *žoq* appears to be minimal (Muhamedowa, 2016, p. 171).[14] In Yakut, a Turkic language spoken in Siberia that historically originates in the region north of Mongolia, there is a similar difference between a negative suffix *-BEt* (with variants) and a negative existential *suoχ ~ suoʁa* that is cognate with Kazakh *žoq*. While past and present tense are marked with the suffix, e.g. *kel-**bep**-pin* 'come-NEG-1SG.COP', the negative existential is used in the future tense, e.g. *kel-**ie(ʁ)**-(i)m **suoʁa*** 'come-FUT-1SG.POSS NEG' (Ebata, 2011, p. 191). Note the possessive marker, which indicates a nominal rather than a verbal form. See Veselinova (2016, pp. 159–163) for the development in some other Turkic languages.

14. Kazakh also shows the development of the negative copula *emes* to a standard negator that in this case "implies that an event has never happened", *kel-gen emes-pin* 'come-PERF.NEG-1SG' (Muhamedowa, 2016, p. 187). This separates it from the other languages surveyed here, although some Mongolic languages have a similar development (see Brosig, 2015).

The Sinitic language Mandarin has two preverbal standard negators, *bù ~ bú* 不 and *méi* 没. The latter is also an existential and possessive negator that is often combined with the existential or possessive *yǒu* 有 (47a). This somewhat contradicts the negative existential cycle, but need not concern us for the moment. Used as a standard negator, *méi* is used to negate perfective aspect (47b) while *bù ~ bú* is the default negator that is otherwise used (see Wiedenhof, 1994 for details).

Mandarin (own knowledge)

(47) a. 我没有书。
wǒ ***méi.yǒu*** *shū.*
1SG NEX book
'I don't have (a) book(s).'

b. 我没有去。
wǒ ***méi.yǒu*** *qù.*
1SG NEG go
'I did not go (there).'

Mandarin differs from the other languages with respect to word order. In the creolized Sinitic language Wutun (Janhunen et al., 2008, pp. 100f.), there are the two prefixes *be-* for irrealis negation (Mandarin *bù ~ bú*) and *mi-* for negation of the past tense and realis (Mandarin *méi*), e.g. *mi-qe-lio* 'NEG-eat-PRF'. The latter has the form *mi(-yek)* as negative existential (Mandarin *méi-yǒu*). This illustrates the fact that a different source construction leads to a different outcome, even if the semantic shift is identical. These few examples are sufficient to illustrate that the newly formed negative strategy usually has a special semantics (e.g., aspect, tense) and is in competition with the original construction. Most likely this was also the case for Manchuic before the earliest recordings.

5. Conclusion

Northeast Asia is a region of widespread bi- and multilingualism (e.g., Janhunen, 1996, pp. 81–85) and of strong convergence of different languages. In such cases, boundaries between different languages are not absolutely clear-cut. The best proposals for a model needed to capture the linguistic knowledge of multilingual speakers are what has been called a *multilingual* CxG (Hilpert & Östman, 2014, p. 138) or a *diasystematic* CxG (Höder, 2012). This paper has attempted to contribute to our understanding of constructions across languages and has shown that even networks of constructions and instances of constructionalization may be shared by different languages. Nevertheless, I prefer the term *areal construction*

grammar, which is broader in scope, permits the integration of previous research in areal linguistics, and makes clear that multilingualism is not necessary for all areal phenomena (e.g., loanwords). Areal CxG is a useful cover term that is broad enough for a variety of different approaches, which share a similar theoretical background and a focus on language contact phenomena. One such approach is *relabeling* – "a process that consists in assigning a lexical entry a new label derived from a phonetic string drawn from another language." (Lefebvre, 2015, p. 1). A Khamnigan Evenki dialect, for example, has relabeled its negative existential *aacin* on the basis of Khamnigan Mongol *vgvi* (cf. Janhunen, 1991, p. 58). Another such approach presented in this paper is *shared constructionalization* – an analogous creation of a new form-meaning pair of any kind in a group of often geographically contiguous languages due to language contact (cf. Heine & Kuteva, 2010; Aikhenvald, 2013).

Table 7 summarizes and compares all the constructions and constructs discussed in Sections 2 and 3. With Höder (2012) one may call this an *interlingual constructional network* shared by Manchu and Mongolian.

Several changes in Manchu and Mongolian certainly qualify as an instance of shared constructionalization (both lexical and grammatical) with similar subsequent constructional changes. From the perspective of CxG, the development of the new standard negative construction may be characterized as an extension of the interlingual constructional network shared by these languages to form a new type node with new morphosyntax and new meaning. The change of meaning is from a partly contentful meaning (negative existence) to a more procedural one (standard negation) (e.g., Trousdale, 2014). The exact meaning of negation as such is still a matter of debate, but it certainly affects *how* we construe (or simulate) a certain situation rather than *what* we construe (Bergen, 2012, pp. 53ff., 108, 140ff.). Negative existence, on the other hand, has an additional semantic component that can be directly observed in analyzable existential negators such as Italian *non c'è* 'NEG EX.SG'. To what degree the new constructions still evoke the source network is an open question that can only be settled with the help of sophisticated experiments that go beyond the possibilities of this paper. But in any case, the original construction as well as its network are still intact and exist alongside the new one. All languages addressed in this paper share the development from negative existential to a standard negator and thus the development towards a more procedural meaning. Several details differ from language to language, but apart from Sinitic all languages share a similar syntactic behavior of the negative existential and most languages require a nominalization of the verb. Furthermore, all the languages still preserve another negator. However, as shown by Veselinova (2016) this does not necessarily indicate a recent development as intermediate stages on the negative existential cycle may be relatively stable. The survey of lan-

Table 7. A comparison of constructions and constructs in Manchu and Mongolian

Meaning	Manchu	Mongolian
negative existence	[N *aq^{h}ʊ*]	[N *=güi/ügüi*]
'without horse'	*morin akū*	*mory=güi*
'without face > shameless'	*derakū* < *dere akū*	*nüür=güi*
'without matter > no problem'	*baitakū* < *baita akū*	(T. Dagur *bajt uwee*) (*yum=güi* 'nothing')
'to become non-existent > to die'	*akū o-*	*ügüi bol-*
negative possession or existence	[N_X *-de* N_Y *aq^{h}ʊ*], (Aihui M. ***bi.dzə**-qo*)	[N_X *-d/t* N_Y ***bai-*** PTCP *=güi*]
attributive privative construction	[[N *aq^{h}ʊ*] ATTR N]	[[N *=güi*] ATTR N]
'name NEX finger > ring finger'	*gebu akū simhun*	*ner=güi xuruu*
negative copula for adjectives	[ADJ *aq^{h}ʊ*]	[ADJ *=güi/ügüi*]
'not good'	*sain akū*	*sain ügüi*
'not feeling well, …'	*elhekū* < *elhe akū*	*enx=güi*
standard negation	[V-PTCP *aq^{h}ʊ* (AUX-)], (*-rakū* > Sibe *-qu*)	[V-PTCP *=güi* (AUX-)] (Dagur *-gw uwei-* > *-wwei*)
comparative construction	[N_X N_Y *-de isi-raq^{h}ʊ*]	[N_X N_Y *-d/t xür-* PTCP *=güi*]
negative obligation 'must not'	[V-*ci o.jo-raq^{h}ʊ*]	[V-*wel bol-* PTCP *=güi*]
negative one word answer 'no'	*akū*	*ügüi*
negative alternative question	*akūn?* < *akū* + *=ni*	*ügüi=y.uu?*

guages was not intended to be exhaustive and the exact extent of the area as well as details of individual languages have yet to be investigated. Especially problematic are the questions whether the cause of the development in at least some of the languages is just a parallel development (e.g., in Dunan) and in which direction possible language contact proceeded (e.g., Thomason, 2010). However, the overall distribution of the languages investigated suggests an origin in Mongolia, Manchuria, or China. Independent of the question in which language the change occurred first, there must have been a cause that led to the change in the first place. A possible reason may be the need for emphatic negation that could only be expressed with a new construction (Croft, 1991). An argument by Janhunen (2012a, p. 253) for Mongolian is basically also valid for Jurchenic. As we have seen, both Tungusic and Mongolic had preverbal negators and Janhunen argues that

this is an atypical syntactic position for these languages. In fact, languages with SV word order such as Mongolian or Manchu tend to have VNeg word order as well (Dryer, 2013, p. 288).

Special abbreviations

. old or irrelevant morpheme boundary
NEX negative existential/existence
PN personal/proper/place name

References

Aikhenvald, A. Y. (2013). Areal diffusion and parallelism in drift: Shared grammaticalization patterns. In M. Robbeets, & H. Cuyckens (Eds.), *Shared grammaticalization. With a special focus on the Transeurasian languages* (pp. 23–41). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.132.07aik

Aixinjueluo, Y. (1990). Tantan manyude jingyu 5 [Discussing the Jing dialect of Manchu]. *Manchu Studies*, 11, 22–37.

Aixinjueluo, Y. (1991). Tantan manyude jingyu 6 [Discussing the Jing dialect of Manchu]. *Manchu Studies*, 13, 3–16.

Barðdal, J., & Gildea, S. (2015). Diachronic construction grammar. Epistemological context, basic assumptions and historical implications. In J. Barðdal, E. Smirnova, L. Sommerer, & S. Gildea (Eds.), *Diachronic construction grammar* (pp. 1–49). Amsterdam: John Benjamins. https://doi.org/10.1075/cal.18.01bar

Bergen, B. (2012). *Louder than words. The new science of how the mind makes meaning*. New York: Basic Books.

Birtalan, Á. (2003). Oirat. In J. Janhunen (Ed.), *The Mongolic languages* (pp. 210–228). London: Routledge.

Brosig, B. (2015). Negation in Mongolic. *Journal de la Société Finno-Ougrienne*, 95, 67–136.

Chaoke, D. O. (2006). Xiandai xiboyu kouyu yanjiu [*Research on modern spoken Sibe*]. Peking: Renmin chubanshe.

Chaoke, D. O. (2007). Nanmu elunchunyu yanjiu [*Research on the Nanmu dialect of Oroqen*]. Peking: Minzu chubanshe.

Chaoke, D. O. (2009). Ewenkeyu cankao yufa [*Reference grammar of Solon*]. Peking: Zhongguo shehui kexue chubanshe.

Croft, W. (1991). The Evolution of negation. *Journal of Linguistics*, 27(1), 1–27. https://doi.org/10.1017/S0022226700012391

Di Cosmo, N. (2006). *The diary of a Manchu soldier in seventeenth-century China. "My service in the army", by Dzengšeo*. London: Routledge.

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, 17(4), 463–489. https://doi.org/10.1515/COG.2006.015

Dryer, M. (2013). On the six-way word order typology, again. *Studies in Language*, 37(2), 267–301. https://doi.org/10.1075/sl.37.2.02dry

Ebata, F. (2011). Sakha (Yakut). In Y. Yamakoshi (Ed.), *Grammatical sketches from the field* (pp. 179–212). Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Fried, M. (2013). Principles of constructional change. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 419–437). Oxford: Oxford University Press.

Georg, S. (2003). Ordos. In J. Janhunen (Ed.), *The Mongolic Languages* (pp. 193–209). London: Routledge.

Gorelova, L. M. (2002). *Manchu grammar*. Leiden: Brill.

Hauer, E. (2007). *Handwörterbuch der Mandschusprache*. Second edition. Wiesbaden: Harrassowitz.

Heine, B. & Kuteva, T. (2005). Language contact and grammatical change. Cambridge: Cambridge University Press.

Heine, B., & Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (Ed.) *Handbook of language contact* (pp. 86–105). Malden: Wiley-Blackwell. https://doi.org/10.1002/9781444318159.ch4

Hilpert, M., & Östman, J.-O. (2014). Reflections on constructions across grammars. *Constructions and Frames*, 6(2), 137–142. https://doi.org/10.1075/cf.6.2.01int

Höder, S. (2012). Multilingual constructions. A diasystematic approach to common structures. In K. Braunmüller, & C. Gabriel (Eds.), *Multilingual individuals and multilingual societies* (pp. 241–257). Amsterdam: John Benjamins. https://doi.org/10.1075/hsm.13.17hod

Hölzl, A. (2015). A typology of negation in Tungusic. *Studies in Language*, 39(1), 118–158. https://doi.org/10.1075/sl.39.1.05hoe

Hölzl, A. (2017). New evidence on Para-Mongolic numerals. *Journal de la Société Finno-Ougrienne*, 96, 21–37.

Ikegami, J. (1999 [1979]). Manshuugo to tsunguusugo. Sono koozoojoo no sooiten to mookogo no eikyoo. In J. Ikegami (Ed.), *Manshuugo kenkyuu* (pp. 344–358). Tokyo: Kyuuko Shoin.

Janhunen, J. (1991). *Material on Manchurian Khamnigan Evenki*. Helsinki: University of Helsinki.

Janhunen, J. (1996). *Manchuria. An ethnic history*. Helsinki: The Finno-Ugrian Society.

Janhunen, J. (2003a). Proto-Mongolic. In J. Janhunen (Ed.), *The Mongolic languages* (pp. 1–29). London: Routledge.

Janhunen, J. (2003b). Khamnigan Mongol. In J. Janhunen (Ed.), *The Mongolic languages* (pp. 83–101). London: Routledge.

Janhunen, J. (2006). The Mongolic languages. In K. Brown (Ed.), *Encyclopedia of language & linguistics*, Second edition (pp. 231–234). Oxford: Elsevier. https://doi.org/10.1016/B0-08-044854-2/02095-2

Janhunen, J. (2012a). *Mongolian*. Amsterdam: John Benjamins. https://doi.org/10.1075/loall.19

Janhunen, J. (2012b). The expansion of Tungusic as an ethnic and linguistic process. In A. Malchokov, & L. J. Whaley (Eds.), *Recent advances in Tungusic linguistics* (pp. 5–16). Wiesbaden: Harrassowitz.

Janhunen, J. (2012c). Khitan. Understanding the language behind the scripts. *SCRIPTA*, 4, 107–132.

Janhunen, J., Peltomaa, M., Sandman, E., & Dongzhu, X. (2008). *Wutun*. München: Lincom Europa.

Kałużyński, S. (1977). *Die Sprache des mandschurischen Stammes Sibe aus der Gegend von Kuldscha*, Volume 1. Warschau: Państwowe Wydawnictwo Naukowe.
Kane, D. (1989). *The Sino-Jurchen vocabulary of the Bureau of Interpreters*. Bloomington: Indiana University Press.
Kim, J., Ko, D., Chaoke, D.O., Han, Y., Piao, L., & Boldyrev, B.V. (2008). *Materials of spoken Manchu*. Seoul: Seoul National University Press.
Kiyose, G.N. (1977). *A study of the Jurchen language and script: Reconstruction and decipherment*. Kyoto: Shōwado.
Kiyose, G.N. (2000). Genealogical relationship of Jurchen dialects and literary Manchu. *Central Asiatic Journal*, 44(2), 177–189.
Langacker, R.W. (2008). *Cognitive Grammar. A basic introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331967.001.0001
Lefebvre, Claire. (2015). *Relabeling in language genesis*. Oxford: Oxford University Press.
Malchukov, A.L. (2013). Verbalization and insubordination in Siberian languages. In M. Robbeets, & H. Cuyckens (Eds.), *Shared grammaticalization. With special reference on the Transeurasian languages* (pp. 177–208). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.132.14mal
Mu Yejun. (1986). Alechuka manyu de shuci yu gezhuci [Numerals and case markers in Alchuka Manchu]. *Manchu Studies*, 2, 2–17.
Mu Yejun. (1987). Balayu [The Bala language]. *Manchu Studies*, 2, 2–31, 128.
Mu Yejun. (1988). Lun balayu de yuyin bianhua [Discussing the sound changes of Bala]. *Manchu Studies*, 6, 1–26.
Muhamedowa, R. (2016). *Kazakh. A comprehensive grammar*. London: Routledge.
Nedjalkov, I. (1994). Evenki. In P. Kahrel & R. van den Berg (Eds.), *Typological studies in negation*, 1-34. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.29.02ned
Nedjalkov, V.P., & Otaina, G.A. (2013). *A syntax of the Nivkh language. The Amur dialect*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.139
Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language*, 14(2), 77–202. https://doi.org/10.1075/fol.14.2.04noe
Norman, J. (2013). *A comprehensive Manchu-English dictionary*. Cambridge: Harvard University Asia Center.
Robbeets, M. (2014). The development of negation in the Transeurasian languages. In P. Suihkonen, & L.J. Whaley (Eds.), *On diversity and complexity of languages spoken in Europe and North and Central Asia* (pp. 401–420). Amsterdam: John Benjamins.
Róna-Tas, A. (2017). Khitan Studies I. The graphs of the Khitan small script 2. The vowels. *Acta Orientalia Academiae Scientiarum Hungaricae*, 70(2), 135–188.
Shapiro, R. (2010). Chinese Pidgin Russian. *Journal of Pidgin and Creole Languages*, 25(1), 5–62. https://doi.org/10.1075/jpcl.25.1.02sha
Sinor, D. (1949). Le verbe mandjou. *Bulletin de la Société de Linguistique de Paris*, 45, 146–155.
Skribnik, E. (2003). Buryat. In Juha Janhunen (Ed.), *The Mongolic Languages* (pp. 102–128). London: Routledge.
Stary, G. (2009). *Geister, Dämonen und seltsame Tiere. Ein mandschurisches Liaozhai zhiyi aus Xinjiang*. Wiesbaden: Harrassowitz.
Steen, F., & Turner, M. (2013). Multimodal Construction Grammar. In M. Borkent, B. Dancygier, & J. Hinnell (Eds.), *Language and the Creative Mind* (pp. 255–274). Chicago: University of Chicago Press.

Thomason, S. (2010). Contact explanations in linguistics. In R. Hickey (Ed.) *Handbook of language contact* (pp. 31–47). Malden: Wiley-Blackwell. https://doi.org/10.1002/9781444318159.ch1

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional change*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Trousdale, G. (2014). On the relationship between grammaticalization and constructionalization. *Folia Linguistica*, 48(2), 557–577. https://doi.org/10.1515/flin.2014.018

Tsumagari T. (2003). Dagur. In J. Janhunen (Ed.), *The Mongolic Languages* (pp. 129–153). London: Routledge.

Tsumagari T. (2009). Grammatical outline of Uilta (revised). *Journal of the Graduate School of Letters*, 4, 1–21.

Verbiest, F. (1682). *Elementa linguae tartaricae*. Paris: Moette.

Veselinova, L. (2014). The negative existential cycle revisited. *Linguistics*, 52(6), 1327–1389.

Veselinova, L. (with Skirgård, H.). (2015). Special negators in the Uralic languages: Synchrony, diachrony and interaction with standard negation. In M. Miestamo, A. Tamm, & B. Wagner-Nagy (Eds.), *Negation in Uralic languages* (pp. 547–599). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.108.20ves

Veselinova, L. (2016). The negative existential cycle through the lens of comparative data. In E. van Gelderen (Ed.), *Cyclical change continued* (pp. 139–188). Amsterdam: John Benjamins.

von Möllendorff, P. G. (1892). *A Manchu grammar with analysed texts*. Shanghai: American Presbyterian Mission Press.

Wagner-Nagy, B. (2015). Negation in Selkup. In M. Miestamo, A. Tamm, & B. Wagner-Nagy (Eds.), *Negation in Uralic languages* (pp. 133–158). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.108.05wag

Wang Q. (2005). Manyu yanjiu [*Research on Manchu*]. Peking: Minzu chubanshe.

Wiedenhof, J. (1994). Standard Mandarin. In P. Kahrel, & R. van den Berg (Eds.), *Typological studies in negation* (pp. 93–124). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.29.06wie

Wuge S., & Cheng, M. (1730). Manju nikan hergen-i Cing wen ki meng bithe. Man han zi Qingwen qimeng [*Enlightenment on the Manchu language in Manchu and Chinese scripts*], Volume 3. http://gallica.bnf.fr/. Accessed 22 April 2016.

Yamada, M, Pellard, T., & Shimoji, M. (2015). Dunan grammar (Yonaguni Ryukyuan). In H. Patrick, S. Miyara, & M. Shimoji (Eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (pp. 449–478). Berlin: De Gruyter Mouton.

Yamakoshi, Y. (2007). Shinehen buriyaatogo no clitic [The clitics in Shineken Buryat]. *Asian and African Languages and Linguistics*, 2, 1–15.

Yamakoshi, Y. (2011). Shineken Buryat. In Y. Yamakoshi (Ed.), *Grammatical sketches from the field* (pp. 137–177). Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Yu W., Kwon, J., Choi, M., Shin, Y., Borjigin, B., & Bold, L. (2008). *A study of the Tacheng dialect of the Dagur language*. Seoul: Seoul National University Press.

Zhang P. (2013). *The Kilen language of Manchuria. Grammar of a moribund Tungusic language*. Hong Kong: University of Hongkong. (Doctoral dissertation.)

Zhao J. (1989). Xiandai manyu yanjiu [*Research on modern Manchu*]. Peking: Minzu chubanshe.

Zikmundová, V. (2013). *Spoken Sibe, Morphology of the inflected parts of speech*. Prague: Karolinum.

A Radical Construction Grammar approach to construction split in the diachrony of the spatial particles of Ancient Greek

Some theoretical preliminaries

Emanuel Karlsson
Uppsala University

Within the context of a Radical Construction Grammar approach to grammaticalization, this paper presents *construction split* as a model of the diachrony of *spatial particle* constructions from Proto-Indo-European to Ancient Greek, accounting for the behaviour of Greek preposition and prefix constructions. The paper discusses some theoretical preliminaries for the RCG approach and for long-term construction developments, with special reference to phenomena in Ancient Greek. An argument is made for central tenets of RCG: there is no structure external to language-specific constructions; language change is fundamentally a matter of degree and ultimately dependent on the pragmatics of discourse; constructions are defined in terms of *cues*. This is framed in a conceptual discussion on the complementarity of grammaticalization and construction perspectives.

Keywords: Radical Construction Grammar, grammaticalization, utterance selection, construction split, Articulatory Phonology, Ancient Greek, spatial particles

1. Introduction

This paper presents a *Radical Construction Grammar* (RCG; Croft, 2001) diachronic analysis of a group of morphemes in Ancient Greek that will be referred to collectively as *spatial particles*. Etymologically, these are space words forming a class of particles in Proto-Indo-European (PIE). By the time of Classical Greek (5th century BCE), these particles have developed cognate constructions that are partially contrastive. For example, the sets of cognates in (1) and (2),

https://doi.org/10.1075/cal.21.c10

respectively the spatial particles *kata* and *epi*, are conventionally described in (1a, 2a) as prepositions and in (1b, 2b) as verbal prefixes.

(1) a. Classical Greek, Hdt. 1.75
tò stratópedon hidryménon *katà* nṓtou
the.ACC.SG camp.ACC.SG fortified.ACC.SG *down* back.GEN.SG
láboi
take.AORIST.OPT.3SG
'it would overtake the fortified camp along the back'

b. Classical Greek, Hdt. 5.71
kata-labeîn tḕn akrópolin epeirḗthē
down-take.AORIST.INF the.ACC.SG citadel.ACC.SG try.AORIST.IND.3SG
'he attempted to take down (=capture) the citadel'

(2) a. Classical Greek, Hdt. 7.54
epì térmasi toîsi ekeínēs génētai
to border.DAT.PL the.DAT.PL that.GEN.SG become.AORIST.SUBJ.3SG
'he would come to its borders'

b. Classical Greek, Hdt. 7.34
epi-genómenos kheimṑn
on-become.AORIST.PTCP.NOM.SG storm.NOM.SG
'a storm having come up / upon [them]'

This situation, as in many other IE languages, is partly parallel to the contrast between verbal particles and prepositions found in English, as in the translations for (1–2).

In earlier stages of Ancient Greek (viz., Homeric Greek, circa 7th century BCE), such a clear distinction between two constructions is still in the process of forming. This paper argues that these patterns have been pulling apart continuously out of a single construction since PIE; this development can be described as a construction *splitting* into multiple descendent constructions through multidirectional grammaticalization processes. It is also suggested that senses of individual particles develop in different directions under the influence of the constructions in which they are used; thus the evolving function of a construction is a driving force in the change of the lexical semantics contributed by individual morphemes in construction slots.

Taking an RCG perspective, this paper explores some theoretical implications of comparing cognate constructions. It attempts a synthesis of theoretical components relevant to construction diachrony and grammaticalization in an RCG model of construction formation. It is argued that construction split and grammaticalization in general are best modelled using *exemplar theory* (recently, Bybee, 2010), ultimately in terms of pragmatic mechanisms where phonology may

constitute a key component; this is a consequence of RCG being a cue-based approach to constructions. The Ancient Greek spatial particle constructions are conducive because their development can be traced over a vast time-span. This paper will focus on the theoretical preliminaries for the analysis of this development, in keeping with the topic of this volume.

The theoretical argument for RCG pursued here sees pragmatics as the fundamental mechanism in language change. In the philosophy of language, the *pragmaticist* position holds that meaning is created in discourse and that language(s) continually take(s) shape in the course of cumulative practice. As a consequence, there is no external–universal frame-of-reference against which meaning, or language structure, could be judged. Pragmaticist views have been articulated rather intermittently, starting with Peirce and Wittgenstein; among the more recent, one may mention Hirsch (1967, p. 4), Chatterjee (1982), Gumperz (1982, pp. 205ff), Lakoff (1987, pp. 111ff), Givón (1989), Croft (2000, 2001), Croft & Cruse (2004, pp. 97ff), Bakker (2010, p. 151). Taking such a view, *language* can be said to consist of minimal semiotic exchanges of utterance *cues* (linguistic 'substance' or, in a sense, 'form') for discourse *effects* ('function' or 'meaning' in the broadest sense). The relations between these, though routinized and systematized to various degrees, are necessarily in flux – they are negotiated through constant reapplication to ever-new discourse situations. Following Chatterjee (1982, p. 338), "what we lazily call the *grammar* of a language is an unending dialectic of utterance and understanding in context". Many implications of the pragmaticist approach to modelling language change remain to be explored, not least the relationship of creativity to convention.

Putting pragmaticist ideas into practice, *grammaticalization* conceptualizes language as an ongoing process – *contra* the traditional notion of language as a metaphysical entity with a set of definitional properties (cf. Givón, 1979, 1989, 2001, pp. 4ff, 27ff; Croft, 2000, pp. 2ff, 2001, passim). Accordingly, grammaticalization invites functional–cognitive motivations for language structure: form–function iconicity, conceptual metaphor, pragmatic inference, prototype categorization, spreading activation, etc. Within the grammaticalization research paradigm, this is closely tied to the empirical appreciation of actual discourse, synchrony–diachrony interrelationships, and cross-linguistic diversity.

In equating *constructions* with semiotic units, *construction grammar* is able to provide a unified framework of language structure. This precludes a discrete level of 'grammatical' structure. More comprehensively, and more controversially, this allows Croft (2001, pp. 203ff) to eliminate the concept of 'syntactic relations' and hence any external frame-of-reference for grammatical knowledge such as would be provided by primitive categories. Among other things, this does present a problem for grammaticalization conceived of as a distinct *process* of language

change, because of the difficulty of defining what makes an outcome 'more grammatical'. This difficulty was in fact recognized by Hopper & Traugott (1993, p. 23), who opted to delimit grammaticalization-as-process rather arbitrarily by excluding word order changes. But the issue may be mostly terminological – taking *grammaticalization* as a descriptive term for a subset of construction developments, rather than a process in its own right, does not deny the insights drawn from it (cf. i.a. Fischer, 1999; Bybee, 2010, pp. 112f).

This paper emphasizes the complementarity of the grammaticalization and construction perspectives. The starting point here is the framework presented in Croft's tandem volumes (2000, 2001) on *utterance-based* evolutionary theory and RCG. The idea of grammaticalization is refurbished in Croft's utterance selection model: language traits are created, modified and adopted in individual utterances primarily to produce pragmatic effects, exploiting the indeterminacy of form–function mapping (Croft 2000, p. 118). This paper seeks to problematize the causes and mechanisms of construction formation from an RCG point of view – especially how constructions *split* through the course of divergent grammaticalizing developments pulling in different directions, as is exhibited by the spatial particles. An unprejudiced approach to such changes must begin by dispelling some analytical preconceptions regarding syntactic relations (dependency) and linear order (constituency). Hence the next section begins by emphasizing how RCG, as opposed to traditional views, takes a *cue-based* approach to constructions, and goes on to explore the implications of this for grammatical change. Based in this, Section 3 outlines theoretical concerns for a model of construction split. Section 4 presents an account of the diachrony of the spatial particles.

2. A RCG approach to the units of grammatical change

2.1 Utterances as constructions, constructions as cues

This section critically examines what constitutes the units of grammatical change. These technical issues are all too often passed over in favour of traditional preconceptions, but it is absolutely essential for any account of grammaticalization to (re)consider them. RCG (Croft, 2001, 2005; cf. also Bybee, 2010) recognizes no primitive grammatical categories or 'syntactic relations' because there is no need or unambiguous evidence for them (Croft, 2005, pp. 288f). This means that RCG does not presuppose that 'form' has to be analysed *ab initio* in terms of such categories. The basic unit of language is the construction; overt coding serves to identify constructions and keep track of semantic components (Croft, 2005, pp. 300ff). Any relationship between constituent parts across constructions is an empirical

issue relating to the individual constructions (Croft, 2001, p. 57). The definition of 'form' becomes simply that of overt coding (including the lack of one).

This fundamental realignment away from structuralism seems to have had limited impact. As Holmes & Hudson (2005, p. 244) put it, construction grammar has habitually been "assuming, without discussion, that sentences are built box-wise out of phrases, so the assumed model of sentence structure is a version of phrase structure". A dependency analysis as espoused by Holmes & Hudson (and i.a. Haspelmath, 1998) may be a step in the right direction, but it is not the only alternative to constituency phrase structure. RCG is neither constituency-based nor dependency-based. Instead it offers a third alternative, which may be called the *cue-based* approach to constructions; this is of fundamental importance to the grammaticalization process envisaged by RCG, answering Bybee's call for a revised synchronic model (2010, p. 120, quoted in 3.1). As outlined by Croft (2005):

> cues taken as a whole provide a structural Gestalt which aids the hearer in identifying the construction, and hence its elements and the correspondence relations to its semantic structure (p. 302)
>
> What occurs in natural discourse are constructions, that is, complex syntactic units: we do not hear individual words with category labels attached to them. (p. 283)

Constructions under the RCG definition consist of "properties of the utterance's actual physical (phonetic) form, and as such are available to the hearer without positing any abstract constituent structure" (Croft, 2005, pp. 301f). Hence a cue "does not code relations between syntactic elements" (p. 301). Constituents are not the indispensable elements of 'form', but are – when they are in evidence – derivative of chunking and contiguity-iconic semantics (Bybee, 2010, pp. 136ff; Croft, 2001, pp. 175ff). Indeed, "constituency is gradient and emergent" (Bybee, 2007, pp. 268ff) and "there is no unique constituent structure valid across all constructions in a language" (Croft, 2005, p. 301). The language user does not primarily generate utterances as parsing-trees, but as constructions – and constructions are defined in terms of cues. Utterances may have partial parsing, coexisting parses with varying degrees of activation, or no parsing (cf. also Bybee, 2010, pp. 95f) – ultimately as a matter of context-specific construal. Constructions rely on the mapping of cues to semantics, as played out in discourse interpretation.[1]

1. This, however, differs considerably from the 'emergent grammar' of Hopper (1998), because RCG recognizes *constructions* as synchronically more-or-less stable and rule-bound, in the sense of Givón (2001, pp. 27ff).

It follows that there is also no obligatory 'head–dependent' dichotomy that cross-sects constructions; to assume such an organizational principle is due, in no small part, to the traditional metaphor of some words 'governing' others with regard to some property (like prepositions or verbs the choice of case). In RCG, instead, any notion of 'headhood' resides in semantics: a 'head' is defined as the *principal information-bearing unit* (Croft, 2001, pp. 258f). The only 'relations' are semantic properties of discourse referents, coded in a certain way. And because it is semantic, "in the process of grammaticalization head status gradually shifts" (Croft, 2001, p. 259).

In accounts that assume constituent phrases and/or dependencies to be obligatory, these units are postulated even in the absence of cues – the contention being that utterances must at any time organize unambiguously into covert structures. It is the logical argument of RCG (Croft, 2005, pp. 282f) that there is no unique, non-circular way to assign covert, 'underlying' structure from the 'surface' evidence of overt coding: rather, accounts that make such assumptions have matched *some* overt coding *ad hoc* to preconceived notions of covert structure (cf. Croft, 2001, pp. 41ff; 2005, p. 281; Bybee, 2010, pp. 120ff).

2.2 Deconstructing syntactic relations: Adjacency, agreement, linear order

RCG deconstructs syntactic relations – the traditional primitives – into component cues, these being the units of grammatical change. Thus there are various cues that identify 'elements taken together' for semantic mapping. The most obvious examples are agreement markers and adjacency – both are iconic cues, but independently so. They need not match. Indeed, agreement in adjacency is best seen as double marking. Judging from languages where agreement is marked only when elements are not adjacent (Croft, 2001, pp. 186f), one may infer that this is where such marking first arises.

Treating both these cues as a single unit of grammatical change makes little sense and requires rampant *ad hoc* fixes. This is of special relevance to the spatial particles. First, particle constructions are ultimately independent of linear order. Second, for an analysis to be able to acknowledge adjacency as being emergent and the tightness preconditioned by adjacency as being a matter of degree, adjacency cannot be the basis for primitive units.

Example (3) is a simple variant of an extremely common pattern in Ancient Greek which exhibits such mismatch of agreement and adjacency. Croft (2001, pp. 186f) notes comparable patterns cross-linguistically.

(3) Homeric Greek, Od. 1.325–7

aoid-òs $_i$	áeid-e	periklyt-ós $_i$ […]	nóst-on $_j$	áeid-e	lygr-ón $_j$
singer-	sing-	well.known-	homecoming-	sing-	arduous-
NOM.MASC.SG	3SG.IMPF	NOM.MASC.SG	ACC.MASC.SG	3SG.IMPF	ACC.MASC.SG

"The famous bard was singing. […] He sang of the arduous journey home."

Here, agreement suffixes cue what pieces of information 'belong together', ignorant of adjacency. In the face of this, the notion of obligatory constituency, as equated with 'bracketing', is either a derailed graphical metaphor (for brackets enclose a contiguous unit) or it is simply biased for adjacency-heavy languages (cf. Givón, 2001, pp. 279ff, contra 'non-configuration'). The activation of cues – the agreeing morphemes – that allocate elements to the respective discourse referents is represented in (3) by indexical letters.

Over and above the agreement constructions, there are 'argument' constructions in (3) that map discourse referents via case cues to frame referents that are assigned by the event construal of the lexical verb.[2] The linear order, finally, is set by the stylistic construction schematized in (4): two different but agreeing nominals (either of them an adjective) referring to the same discourse referent.

(4) $_x$ NOMINAL $_i$ + (…) + VERB + (…) + $_y$ NOMINAL $_i$

The agreeing noun-and-adjective co-reference units in (3) have no 'syntactic' cues of embedding internal structure. There are, simply, co-indexed elements; any nesting relationship is semantic (cf. Croft, 2001, pp. 186f). There is morphological dependency (Croft's *coded dependency*), since adjective asymmetrically agrees with noun gender. But, first, that is a *functional* cue of the adjective construction specifically. Second, the relation involves the specific property 'gender'; it is not a

2. Note that it is assumed here that, at least for complex morphologies (unlike, implicitly for English, Croft, 2001, pp. 53ff, 199), inflections are morphological 'categories' in the sense that they cluster several exponents, include non-concatenative cues, and map non-uniquely to semantic complexes, linking forms across paradigms. Many-to-many 'cluster constructions' crucially allow for inflections to intersect. E.g., the Greek *nominative* constitutes a single functional inflection but is differently cued across the genders, declensions and numbers. Further, Ancient Greek forms like *títhēmi* 'put' and *plássō* 'shape' may constitute the same inflectional complex, 1SG.PRESENT.INDICATIVE.ACTIVE, while having no phonetic cue in common. Inflections cue slots and produce the respective forms of lexical items (stored as networks of forms). Note there is no such thing as an 'uninflected' verbal or nominal form in Greek, such as would be the case in more agglutinative languages. Ultimately, inflections as units may themselves become associated with emergent meaning (through *categorization*, Croft & Cruse, 2004, pp. 54f), which may in turn affect how verbs assign them for slots.

general, abstract property 'dependence' between elements (cf. 2.1: cue-picking is circular, Croft, 2005, p. 278).

Following Croft (2001, pp. 226ff), all indexical elements refer, and they agree rather with discourse referents. These in turn are established by roles in local constructions, externally by lexical flags like gender, etc. A white horse may be a kind of horse (*profile equivalent*, Croft, 2001, pp. 256ff), but this semantic asymmetry (cf. Croft's *collocational dependency*) need not be explicitly coded. Adjective and noun in (3) simply provide different pieces of information referring to the same discourse referent; the adjectives are not nested in any constructions in which the nouns are not also nested.

Agreement is not absolute and unique, but different elements can agree for different sub-properties. Mismatch or vacillation in agreement patterns is the outcome of different construals of discourse referents and their contextual scope – e.g., group-noun number agreement, the complex interaction of lexical and semantic gender (cf. Croft, 2001, pp. 229ff), or specialized agreement constructions like how Ancient Greek neuter plural obligatorily takes singular verb (cf. Croft, 2001, pp. 231f on Arabic). Relative pronouns typically agree with antecedent for gender–number but take their case as an argument within the relative clause construction (cf. Croft, 2001, pp. 199f). But in Ancient Greek they may alternatively agree with antecedent case, as in (5), suggesting competing motivations. In such cross-clausal agreement, the discourse referent is still construed in the new clause as being under the scope of the event semantics that is being signalled by the antecedent's case. Similarly, in (6), the predicative noun may agree with the discourse referent either in the role of the finite verb or be marked accusative as part of the infinitive construction.

(5) Classical Greek, Xen. An. 1.3.17
tō̃i$_i$ hēgemóni$_i$ hō̃i$_i$ (/hòn$_j$) doíē hépesthai
the.DAT.SG leader.DAT.SG REL.DAT.SG (REL.ACC.SG) give.3SG.AOR.OPT follow:INF
'to follow the leader whom he would give'

(6) Classical Greek, Xen. An. 7.1.21
nỹn soi$_i$ éksestin andrì$_i$ (/ándra$_j$) genésthai$_j$
now you.DAT.SG is.allowed.3SG.PRES man.DAT.SG (man.ACC.SG) become:INF
'Now it is possible for you to become a man.'

All this would call, at least, for multiple, sub-element dependencies. Also, it makes little sense to assign the element *hō̃i* in (5) to either of the putative constituent clauses (or mutually exclusive dependencies): it is cued simultaneously as agreeing with *hēgemóni*, this being the dative of *hépesthai*, and as part of the relative con-

struction with *doíē*. Again, picking one of these cues as conclusive for the analysis would be circular.

Cross-linguistically, intersecting and mismatching multiple dependencies may be even more pervasive and obligatory. For starters, languages like Basque and Georgian consistently code more than one referent in the verb, as in (7) and (8) below.

(7) Basque (adapted from Croft, 2001, p. 253)
erosketa-k […] egi-n d- it- u -t
purchase-ABS.PL do-PERF 3ABS- ABS.PL- AUX -1SG.ERG
'I did the shopping.'

The verb in such languages would be a dependent of up to three arguments simultaneously (or just as well the other way around) and phrase structure would be flat, unless some of these cues are disregarded arbitrarily. The point is, these structures are equivalent to cues, no more no less: in RCG, the units of analysis and grammatical change are not covert constituents or relations, but cues.

More complex examples of mismatch noted by Croft include verbs agreeing with a noun's attribute (2005, p. 293) and noun case patterning ergatively while verbal agreement patterns accusatively (2001, p. 153). The latter type of crisscrossing dependency is found spectacularly in one group of Georgian verbs: there is one set of verbal affixes to mark an agent-like role and another set to mark a patient-like role across both the present and the aorist, but case affixes mark these roles differently in the two tenses (cf. e.g. Vamling, 1989, pp. 17ff). This is given as Table 1 and seen in Examples (8–10). (Note that word order is independent of argument structure, that the presence of a pronoun is emphatic, and that the 1SG pronoun is syncretic.)

(8) (is $_i$) m $_j$ -xat'av-s $_i$ (me $_j$)
3SG.NOM $_i$ **1SG.M** $_j$ -paint-PRES.3SG $_i$ 1SG.DAT $_j$
'He $_i$ paints me $_j$.'

(9) (me $_i$) c'eril-s $_j$ v $_i$ -c'er-Ø
1SG.NOM $_i$ letter- **DAT** $_j$ **1SG.V** $_i$ -write-PRES(3SG)
'I $_i$ write a letter $_j$.'

(10) (me $_i$) c'eril-i $_j$ da-v $_i$ -c'er-e
1SG.ERG $_i$ letter- **NOM** $_j$ AOR-**1SG.V** $_i$ -write-AOR(3SG)
'I $_i$ wrote a letter $_j$.'

The two systems of dependencies are independent of one another. Only one of them varies as they intersect with other properties, as with tense here (this may be compared with 'semi-deponent' verbs, ubiquitous in Ancient Greek). This is

Table 1. Case and set of verbal affixes marking referents in one group of Georgian verbs

	Agent-like role	**Patient-like role**
Present	nom., *v*-set	dat., *m*-set
Aorist	erg., *v*-set	nom., *m*-set

straightforward in terms of cues: ergative marking is one of the cues that signal the aorist construction; the ergative also cues the agent-like role; this role is also signalled by the verbal affix (regardless of tense). The pattern originates in diachrony, but it carries synchronic function too: an aorist (completive) has been successfully *accomplished* by the agent, so the patient is stative-like. Thus Georgian case is functionally independent of argument structure as construed by the verbal marking – thus case can serve as a cue for the referent's position on a cline of successful affecter:[3]

(11) ergative > nominative > dative

The alternation in (12) shows the potential functional contribution of the ergative case, irrespective of what the verb construes as argument.

(12) Georgian (adapted from Harris, 1981, pp. 246, 305)
macivri-dan c'qal-i$_i$ (/ʔc'qal-ma$_i$) i-c'vet-a$_i$
refrigerator-from water-**NOM**$_i$ (/ʔwater-**ERG**$_i$) AOR-drip-AOR.3SG.3SG$_i$
Water dripped (ʔon purpose) from the refrigerator.

The point is that an ergative in (12) is odd not because cues would mismatch, but because the contributed semantics (water having agency) conflicts with the speaker's conception of the world.

Crucially, such independence of case and verbal marking stands in stark opposition to the strict dependence of nominative and verbal agreement in most Indo-European languages (including Greek), a situation which likely is what underlies the traditional linguistic conceptions of 'subject' and 'verb phrase', wherein the complex nominative-cum-agreement argument traditionally 'governs' the verb and the verb its 'object'. Which frame referent will be construed as the nominative-cum-agreement ultimately resides in the event construal of the individual lexical verb – but it is traditionally imagined as surfacing a putative primitive entity 'subject' (cf. Croft, 2001, pp. 132ff). Again, once case and verbal marking have been disentangled, argument structures cannot be the fundamental units of grammatical change.

3. Georgian perfect shows yet another constellation: it patterns diachronically with experience verbs, with dat.+*m*-set for ~experiencer and nom.+*v*-set for ~theme. Cf., cross-linguistically, perfects taking dative agents.

While there may be cues that signal constituent parts and relations, these are construction-specific. Morphological bracketing, as in double agreement markers, is an example of an iconic cue that nests units. The Old Georgian case morphology in (13) clearly marks nested phrases. But such double, 'phrase final' relational marking on nouns (again showing multiple dependencies) is very uncommon cross-linguistically.

(13) Old Georgian, Matthew 13.11
c'nob-ad saidumlo-**y** $_i$ sasupevel-**isa** $_j$ ca-ta-**ysa** $_j$ -**y** $_i$
knowing-ADV secret- **NOM** $_i$ kingdom- **GEN** $_j$ heaven-OBL.PL- **GEN** $_j$ - **NOM** $_i$
'to know the secret of the kingdom of the heavens'

When such nesting cues are absent, it seems unwarranted to postulate that such structures must be there covertly. The clitic English genitive, articles and prepositions are not so different when exterior to an attribute, e.g. *a white horse*. Such brackets cue the extent of a phrase. But again, in Ancient Greek, comparable clitics are independent of adjacency, as in (14–15). In (14), neither preposition nor article bracket a contiguous unit, and the article is not contiguous with its would-be head – hence they do not cue constituency. In (15), the article is adjacent to the head, but the attribute is non-contiguous.

(14) Classical Greek, Xen. An. 6.4.19
epì= skēnèn $_i$ ióntes $_j$ tèn $_i$ =
to= tent.ACC.SG.FEM going.NOM.PL.MASC ART.ACC.SG.FEM=
Ksenophō̃ntos $_k$
Xenophon.GEN.SG.MASC
'(they $_j$ are) going to the $_i$ tent $_i$ of Xenophon $_k$'

(15) Classical Greek, Hdt. 7.46
glykỳn $_i$ geúsās $_j$ tòn $_i$ = aiō̃na $_i$
sweet.ACC.SG give.taste.AOR.PTCP.NOM.SG ART.ACC.SG= lifetime.ACC.SG
'having given a taste of the sweet life'

Cues of togetherness do not always line up: there is mismatch, intersection of different patterns, and fluidity of construal. In (15), *glykỳn* and *tòn aiō̃na* are a unit morphologically, but separate linear-order units. Rather than monolithic syntactic units, there are conglomerates of cued properties. What elements do belong together is a real indeterminacy of categorization that speakers face (cf. Croft, 2001, pp. 52, 118ff; Givón, 1989, 2001, p. 6; Bybee, 2010, p. 53), ultimately resolved through pragmatic interpretation.

Elements in Ancient Greek *can* be linearly ordered, as in the stylistic construction in (4), which can also be applied to (15). But any constructional slot

cannot fundamentally be defined in terms of linear order, but simply as being cued in some way in order to allocate elements to discourse referents (cf. Croft, 2001, pp. 229ff); this may be purely morphological, as in case constructions, which, in turn, may appear across multiple linear-order slots.[4]

A major aspiration of RCG is to disentangle syntactic analysis from linear order (Croft, 2001, pp. 175ff). This becomes crucial not only for flexible word order languages like Ancient Greek, but for the treatment of grammaticalization in general: it cannot be assumed that the units of grammatical change work on covert analyses based in linear order. It is especially pertinent to the spatial particles, where both synchronic and diachronic analyses must be able to treat linear ordering as an independent value.

3. Construction split

3.1 Preliminaries

The formation of a new construction through branching, as in *gonna* alongside its source *going to*, was termed *functional split* by Heine & Reh (1984, pp. 57ff, passim), *divergence* by Hopper (Hopper, 1991, p. 22; Hopper & Traugott, 1993, pp. 116ff), *heterosemy* in Lichtenberk (1991). It is given wider treatment in Givón (1989, pp. 54ff, *divergence*), Croft (2000, pp. 118ff, 189, *branching lineage*) and

4. Because of the combinatorial complexity in languages like Ancient Greek, it may be suggested that constructions combining in utterance generation not only *nest* slots, but map slots to one another, possibly via discourse referents. This presupposes some form of accommodation, which would have to be pragmatic in origin; how *multiple inheritance* grammaticalizes seems to have been given short shrift, however. Croft (2001, pp. 25f, 197) refers to it only in passing and vaguely suggests 'partial specification' in the *combining* constructions. The non-contiguous ordering constructions treated by Croft (2001, pp. 187f) are extremely simple (resembling (4)). Somewhat paradoxically, Goldberg (1995, p. 98) instead invokes 'full specification' in the *combined* construction, and seems to dismiss the problem out of hand (similarly in Goldberg, 2006, p. 10; and in Kuningas & Leino, 2006, p. 306). Contrary to what Goldberg asserts, on-line inheritance processes must be the norm in flexible-word-order languages like Ancient Greek, where the number of orders potentially equals the number of possible permutations (not counting clitics), these being not necessarily entrenched and probably not quantifiable (cf. e.g. Denniston, 1952, p. 8; Dover, 1960). An intriguing but ultimately unsuccessful attempt at comprehensive modelling of Ancient Greek is Matić (2003), expanding on Dik (1995). Either word order cannot be regulated by constructions, or constructions are generated on the fly. Certain constructions can be singled out for Ancient Greek, such as (4) – but they include underspecified parts (as (4) does) and their coverage is limited. It may be suggested that flexible word-order is partly a function of *lexical retrieval* (on which cf. Bock, 1982), basically as an extreme case of on-line sequence-iconicity (cf. Givón, 1988, 1989, pp. 107f, 1995,

Bybee (2010, pp. 33ff, *autonomy*). Terminology aside, the latter three accounts are basically compatible as they all explicitly consider the process to be fundamentally one of degree. Specifically – and this importantly echoes the position of Haspelmath (1998) – it is not a matter of abrupt reanalysis, as in a dichotomy of alternative constituent structures, since such 'analysis' presupposes primitive units.[5] Haspelmath argues that the appeal to 'reanalysis' is a result of conceptualizing 'form' as synonymous with constituency (cf. 2.1), where degree has been ruled out *a priori*.

> often the implicit claim seems to be that at some point during the gradual process of grammaticalization, there is an abrupt change in the syntactic structure […] [T]he need for an abrupt structural change is an artefact of a representation in terms of a constituency tree (Haspelmath, 1998, p. 345)

Again, phrase structure becomes an indispensable part of language generation and change only because of the preconception that there must always be a uniquely identifying underlying phrase structure to any utterance (cf. 2.1). Reanalysis in terms of some other existing structure is certainly *one* way in which constructions may change. But the assumption of phrase structure is not made in RCG, and reanalysis is not a necessary component in accounting for changes in form, because there are no *a priori* analytical units. An abrupt reanalysis, whereby some construction replaces one structural patterning with another, cannot be assumed in the absence of clear evidence for analogy to other existing constructions – hence reanalysis cannot be the most fundamental mechanism in language change (*pace* for instance Traugott & Trousdale, 2013, who make the claim that although it appears gradual, change is ultimately a series of discrete microsteps in underlying structure).[6] In an RCG framework, at least, pragmatic ('functional') reinterpretation would be the fundamental mechanism, and it does

---

pp. 51ff). In languages where word order is crucial to assigning roles, lexical retrieval resulting in the wrong order necessitates a repair (Croft, 2001, pp. 190ff discusses some examples of non-contiguity and repair in similar terms); in Ancient Greek there would be no need for repair as long as the case constructions are in place.

5. Note that Croft (2000, p. 120) carefully distinguishes his 'form–function reanalysis' from constituency analysis, and Bybee speaks of "so-called 'reanalysis'" (2010, p. 134) as a matter of convention. Preferrably, the term 'reanalysis' should be abandoned altogether, because it suggests an innatist metalanguage in terms of which any utterance could be *analysed* independently. (This does not preclude 'reanalysis' in the sense of *analogy* to already existing, language-specific constructions.)

6. Again, because *analysis* means taking something in terms of some other structure, and because it is assumed here that universal structures do not exist, reanalysis must be the application to the construction in question of some other structure present in the language in question,

not require discrete structural reanalysis. Further, reanalysis in terms of already existing structures cannot be the *primary* mechanism, for that would reduce the explanation of a trait to an infinite regress – unless, of course, some structures are to be considered universal in the innatist sense. Bybee (2010) concludes:

> The gradual nature of grammaticalization and the lack of evidence for abrupt reanalysis suggest that rather than postulate covert, inherently unobservable changes, we revise our conception of synchronic grammar so that it is more in line with the facts of grammatical change. (p. 120)

Section 2 presented RCG as a synchronic model fulfilling this directive by way of a fundamental reconsideration of the units of grammatical change.

In RCG, the process of construction formation could be termed *construction emergence* or *split*. The latter may be more appropriate when one construction simply branches off into two, as seen below. Bybee, however, only treats this process when it first involves chunking, as in *going to*; the start of the process is then the setting up of a representation of the chunked unit (2010, p. 145), which accrues a more and more distinct function through continuous activation in discourse. Less clearly recognized by Bybee, the future construction proceeding from that chunk involves a *semiotic split* – the original sequence *going to* would inevitably get another round of chunking, but nevertheless it still maintains its old function. Also, there are cases where chunking is not involved simply because there is only one syntagmatic element (though with context).

The completion of a split into two distinct units can be glanced from after-the-fact tests of the ability of the two constructions to occur side-by-side, e.g. *going to go* (cf. Heine & Reh, 1984, pp. 257f; Hopper & Traugott, 1993, pp. 118ff for classifiers, and below for prefix–preposition in Ancient Greek), or not being interchangeable, e.g. **I'm gonna town* is inadmissible (although *going to* can be used for the future tense). Thus the two are no longer perceived as being the 'same', with distinction of form–function. One mechanism for this (as Bybee suggests in passing, 2010, pp. 48ff) is the specification of phonological reduction.

3.2 Some phonological underpinnings of grammaticalization

The most important cues of 'togetherness' is what usually goes under the label 'prosody'. These cues are fundamentally iconic, and more or less contingent on the adjacency of the elements involved, because they rely on the organization of phonological structures (cues) relative to one another. Cf. Croft (2005, p. 301):

---

that is to say analogy to another construction in that language. To posit reanalysis without analogy to existing constructions is to assume universal structures.

> There is physical contiguity of elements, which occurs in greater and lesser degrees of tightness; there are grammatical units defined by their occurrence in intonation units […] and other prosodic properties

This presupposes a continuum of the level of 'phonetic' detail that constructions may specify, making a phonetics–phonology dichotomy pointless. The view of phonology espoused here is compatible with that of Vihman & Croft (2007) and of Bybee (2010, pp. 38ff), but the concern is not primarily with lexical constructions but their interaction with non-lexical constructions.

Cues are a matter of phonological specification (cf. 2.1, Croft, 2005, pp. 301f) and can be defined as properties of phonological *gestures* in the sense of Browman & Goldstein (1992).[7] Construction-specific phonological specifications constitute spans of allowed gestural variation that map to semantic effect. Following *exemplar theory* (i.a. Johnson, 2007), these categorizations directly relate to the span of exemplar tokens in 'phonetic space' – i.e., specific instantiations accumulated through discourse usage, as stored in memory.

Gestural properties may be directly influenced by pragmatic concerns. Relative length and enunciation can be a cue for marking relative saliency of elements in discourse.[8] If gestures are attenuated or amplified, they iconically cue a corresponding pragmatic effect. In Lindblom's terms (1990, Lindblom et al., 1995), there is a continuum between hyperarticulated and hypoarticulated speech, along which an instantiation is proportionately related to its pragmatic effect. Pragmatic closeness of adjacent discourse elements may be cued by the relative contraction of gestures, producing greater gestural overlap. Thus the mechanism for Croft's 'tightness' above translates into the relative 'tightness' of gestures. Such patterns are overlaid on the cues of the constructions encompassed. When these specifications

---

7. The real-time allocation of gestures as cues of constructions can be compared to the *cohort* model of Marslen-Wilson (1987): cues accumulate in a relative activation level of candidates up to a uniqueness point where the construction is identified. As the term implies, 'gestures' are ultimately articulatory. There is, however, evidence (Lindblom, 1990; Houde & Jordan, 2002; Cummins, 2010) to suggest that phonological specifications are stored as goal-oriented 'Gestalts' that integrate perceptual effect. At any rate, given the nature of the articulators, invariant segments are neither *necessary* nor *efficient* for production and perception (cf. also *templates* in Vihman & Croft, 2007, the *groping* effect in Johnson, 2001, and the Firthian term *prosodies*). There *is* some psycholinguistic reality to segments (perhaps downplayed too much in Vihman & Croft, 2007), if only as emergent categorization, at least to the extent that analogies etc. can work on them. Croft (2005, p. 311) is understandably skeptical of "constructions which have formal values but no semantic value", but segments could be conceived as shortcuts having internal structure.

8. Discussed at least as early as Bolinger (1961, pp. 25ff). Cf. i.a. the *Cognitive–Pragmatic Phonetics* of Tatham (1990), Morton (1992), Tatham & Morton (2006).

of form are coupled with a specific function, this coupling may entrench, forming independent constructions. The phonetic traits of reduction pervasively correlate with the pragmatic gradient between content word and function word (Johnson, 2001), thus being a key component in grammaticalization, as in *gonna*.

Marked boundaries between units may thus be signalled by lack of overlap of gestures across the units. Conversely, cued unit boundaries 'fade' as a matter of gradual gestural tightness coupled with increasing pragmatic construal as a single unit under a single semantic scope. This *chunking* is a matter of degree, accumulating with usage (Bybee, 2010, pp. 33ff). The traditional notion of 'clitic' already implies a further distinction than the dichotomy between element and part of element. In terms of the tightness of gestures, cliticness is a matter of degree, along a continuum ranging from stressed focus element to fusional cue.

3.3 A grammaticalization mechanism for split

Haspelmath defines grammaticalization as "the gradual drift in all parts of the grammar toward tighter structures" (1998, p. 318). In 3.2, this notion of 'tightness' (also, Givón, 1979, pp. 207ff, 2001, p. 35; Croft, 2005, p. 301) translated, for form, into gestural organization. On the semantic plane, the mechanism described by Givón (1989, pp. 54ff, 2001, p. 34) as *metaphoric extension of prototype* in category drift is basically compatible with a model of an *exemplar space* of semantic instantiations (also suggested by Bybee 2010, p. 80). A similar view of *exemplar semantics* also appears in Croft (manuscript). This type of change can be conceived as shifts of category spans in exemplar spaces – semantic change is then a continuous process of reconstrual, accumulating with each new discourse instantiation. Cf. Haspelmath (1998):

> grammaticalization fundamentally results from the gradual shift of synchronic variation continua […] A new generation of speakers may take a more grammaticalized form as the basic one […] this form may at some point disappear because a new generation of speakers is no longer exposed to it sufficiently and makes [the more reduced one] their basic form., (p. 345–346)
>
> some elements with an abstract, general meaning prove to be particularly useful in many contexts and increase in frequency. Since ease of perception is less important with frequent items, they are subject to greater phonetic reduction, including merger of adjacent units that commonly occur together. (p. 322)
>
> Expressions that are widely used […] can no longer fulfill the role of a pragmatically salient marker. (p. 318–319)

This is very reminiscent of being a diachronic extension on Lindblom's (1990) Hyperspeech–Hypospeech model of pragmatic variation along a phonetic contin-

uum (cf. 3.2). Thus decreasing duration correlates with predictability and hence informational load (cf. Haspelmath, 2008b; Fenk-Oczlon, 2001; Bybee, 2010, pp. 38ff; and the Gricean maxim of quantity).[9] By extension, semantic 'bleaching' of propositional meaning "goes hand in hand" (Givón, 1979, p. 232) with phonological reduction. Phonetic variations become indexical. An associative iconicity (cf. Fischer, 1999) emerges from the correlation of phonetic modification with what is frequent and predictable – a construction-specific functional cue of increased predictability, hence being expressive of semantic change in grammaticalizing elements. Indeed, it seems warranted to suggest that phonetic reduction is not a byproduct of grammaticalization, but a *mechanism* for signalling elements as less contentful and more 'grammatical'. Cf. Johnson (2001):

> syntactic and semantic context delimits a range of permissible speech reduction patterns. […] function word reductions are 'meaningful' in that grammatical class is signaled by a pattern of variability unique to function words […]
> emphatic pronunciations add a level of gradient salience […] this directly contributes to the listener's on-line construction of pragmatic discourse structure.
> (Johnson, 2001, no pagination)

Pragmatics is fundamentally goal-oriented. Croft (2000, pp. 64ff), in the context of the *utterance selection* model, makes a crucial distinction between *teleological*, as in attributing a goal to an abstract system, and *intentional*, as in speakers' pragmatic goals in discourse contexts. Speakers constantly have motive for being attuned to emphasis and predictability in terms of present discourse – a

9. Note that Bybee earlier considered reduction to be an automatic effect of frequency (i.a. 2004, p. 78, 2007, p. 11; more wavering is Bybee & Hopper, 2001, pp. 11f). That is not phonetically plausible: articulation is a goal-oriented coordination task (Cummins, 2010) and the ability to reach targets *improves* with practice. Articulations do not simply 'wear down' (as Hopper & Traugott, 1993 put it) with extensive use; speakers remain largely (subconsciously) in control of their phonetic variation (cf. Lindblom, 1990). In fact, the difficulty of defining 'ease' of articulation objectively (cf. Ladefoged, 1990) is a problem also for Lindblom's model: the effects of reduction are highly language-specific and the establishment of variation continua seems conventional. At any rate, as also pointed out by Haspelmath (2008b), *predictability* in discourse is fundamental to reduction, not frequency in itself, though the two often align. More recently, Bybee (2010, pp. 37ff) has thoroughly revised her position: reduction occurs *in context* and is not 'automatic'; frequency is not the cause of but only allows for reduction; a more prominent role is given to predictability. One should also note that Haspelmath's (2008a) appeal to frequency–predictability over iconicity concerns specific effects in morphosyntax; as a contribution to the debate which followed Haspelmath's paper, I would suggest Croft's (2008) 'iconicity of distance' may be subsumed under the 'iconicity of sequence' which Haspelmath (2008a) accepts.

mechanism initiating differential correlation of form and function leading to construction split.[10]

A conceptualization of this type of change in terms of exemplar space is illustrated in Figure 1. The spans of the 'cloud' of accumulating exemplar tokens may shift in two dimensions: on the one axis there is a relative differentiation of phonetic form, ultimately a construction-specific modification of gestures; this may correspond on the other axis with a shift in meaning, ultimately as novel discourse construals, via what Givón (1989, pp. 54ff) calls *metaphoric extension* (on terminology, however, cf. Brinton, 1988, pp. 197f). Figure 1 represents schematically such correlation of phonetic and semantic drift.

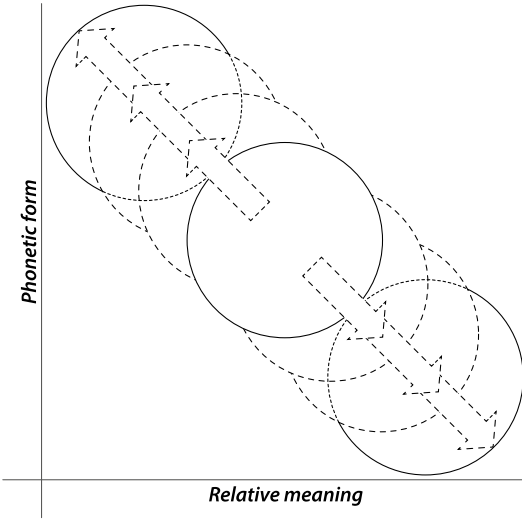


Figure 1. Conceptual model for construction split by metaphoric extension and phonetic differentiation (cf. Figure 4.7 in Heine et al., 1991, p. 112)

When an increasing number of tokens expand these spans of both form and meaning, the result may be construction split. The phonetic differentiation of a group of tokens towards a reduced form like *gonna* (in the direction of relative *hypospeech* to use Lindblom's terminology) is correlated with a gradual change of construal, at first in terms of predictability. Split has occurred when cognate tokens no longer form a contiguous cloud with a single prototype core (cf. the 'bridge' of Givón, 1989, p. 56). From the data of Lorenz (2012, 2013, who calls this

10. Note that this is not a general model for sound change, only one to account for the special case of grammaticalization reduction, as amply attested in examples like *gonna* next to *going to*.

emancipation; cf. also Krug, 2001) it seems that *gonna* has indeed fully split into an autonomous unit: namely, it is defined by its own continuum of variation, showing an independent reduction pattern from that of *going to*. The historical collocation has undergone chunking, followed by a tendency for reduction in pragmatically less weighty and more predictable contexts. This entails a gradual expansion of the phonological span of tokens, correlated with more and more 'grammatical' construal. This emergent distinction becomes the basis for the formation of a new prototype core.

Changes in constructions can be represented on the level of the construction network as the change of form and meaning in existing nodes and the creation of new nodes. Formation of a new node typically implies a differentiation of both form and meaning (cf. Traugott & Trousdale, 2013). Gradual change seems to imply that nodes have a level of activation (cf. Bybee, 2010, p. 47; Croft, 2000, pp. 118f), corresponding to the group of tokens that instantiate them. Within an exemplar framework, node formation can be related to the formation of a new prototype core; increasing recategorization of a new grouping of tokens activates an emergent node. Thus, through metaphoric extension in the course of pragmatic interpretation, inventive tokens gradually cluster to form an expansive group that may 'bud off' from the group of tokens associated with the parent construction, resulting in an independent exemplar cloud with its own prototype core, gradually increasing the relative level of activation of the new construction's node.[11]

3.4 Split of adpositions and affixes

The traditional grammarians' description (for Latin) of preposition as 'governing' case (cf. 2.2) may have been intuited from 'independent' clitic vs. 'bound' inflection. There is a traditional quandary over whether or not adpositions should be heads of eponymous phrases that nest the noun (cf. Croft, 2001, p. 263; Bybee, 2010, pp. 138ff) – a hesitancy which may perhaps be seen in relation to the fact that in many modern Western European languages prepositions no longer 'govern' any case (unlike Latin, and so the origin of the received metaphor becomes opaque).

11. A reviewer notes that change is ultimately a matter of individual speech acts, and so tokens must be discrete and hence change cannot be truly continuous. This again presumes (cf. 2.1, 3.1) that every utterance must be exclusively defined by a uniquely identifying structural analysis. Tokens are indeed discrete events, but this does not imply the either-or discreteness of mental categories applied to those tokens (cf. also Givón, 1989, pp. 38ff). Tokens are positioned along a scale not just phonetically but semantically as well, meaning the relation between the tokens is continuous.

Adpositions basically resemble clitic case markers (cf. Givón, 2001, p. 95; Croft, 2001, pp. 263ff). In case-heavy languages, nouns with adpositions also have case marking: thus in Example (14) in 2.2 above, the preposition *epì* 'to (etc.)' is linked with the accusative. Different cases may further disambiguate different senses; or, as a diachronic remnant, case may be fixed and thus redundant. Since *epì* occurs with all three oblique cases in Ancient Greek, these constitute specific (sub)constructions, each of which is highly polysemous. As all cases also occur with other prepositions, this asymmetry is a many-to-many mapping. In RCG, the combined preposition and case together cue a semantic complex (cf. Croft, 2001, p. 226 for Russian).

A complex interaction of constructions and construals of discourse referents determines semantic interrelationships between verb–adposition–argument. An English preposition *may* be triggered lexically by the verb, even though it is cued as a proclitic on the nominal phrase construction profiling the referent (the 'argument'). Nevertheless, the preposition is making a semantic contribution to the referent. Such indeterminacy of where the preposition 'belongs' is likely linguistically real (cf. Croft, 2000, p. 118). Whether the preposition is more related to the verb or to the argument is ultimately a matter of contextual construal (besides being collocation-specific). Since the preposition refers to one and the same semantic relation, relative to either verb or argument, it will usually be expressed only once as a cue, even though it relates to both.[12] Compare the representation of semantic links in Croft (2005, p. 226), where preposition *together* with case links to the semantic relation between predication and situation.

Besides appearing in constructions with prepositions, Ancient Greek cases may be dependent lexically on the verb, or may be used more 'freely' in adverbial constructions. The difference between an English preposition and an Ancient Greek case construction is rather slight and reflects the diachronic cline between adposition and case marker in relative 'tightness' to the noun (cf. 3.2) and in obligatoriness. Bybee (2010, pp. 138ff) discusses graded analysability and constituency of 'compound' prepositions – arguably, this cline eventually leads to affixation.

The split of the spatial particle construction into preposition and prefix constructions presumably develops through the association and gradual entrenchment of the position of the particle adjacent to the respective pragmatically salient element (*profiling* in Croft's terms; see further 4.1), correlated with gradual tightening of form relative to that element. This split is not abrupt, but is an emergent

12. Such double marking may show up in 'restarts' and the like where verb and argument are separated by a pause. A variant of such double marking phenomena might be what is behind the double particle construction in Ancient Greek (cf. 4.2).

effect of pragmatic construal of relative relationships over cumulative discourse instantiations.

The diachronic development of tighter prefixation may be considered in the light of synchronic phonetic data. In a study of English prefixes, Smith et al. (2012) measured the duration of prefixes in verbs on different points on the cline of lexicalization (entrenchment), and found a consistent difference correlated to the (perceived) contrast between the less lexicalized (e.g., *mistimes*) on the one hand, and the more lexicalized (e.g., *mistakes*) on the other. The more lexicalized the prefix combination, the consistently shorter the duration of the prefix relative to the whole. They conclude:

> reduction in duration [is a cue for distinguishing] the first syllable as either a separate, productive morpheme and hence somewhat independent of the second syllable […], or part of the same morpheme as the stem, and hence less independent.
> (Smith et al., 2012, pp. 700–701)

For the purpose of illustration, a simplified gestural analysis (cf. 3.2; cf. i.a. Marin & Pouplier, 2010) of this effect, building on the data of Smith et al. (2012), is presented in Figure 2, showing gestural contraction as a cue for tighter structures.

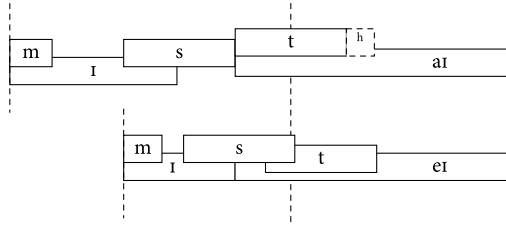


Figure 2. Schematic gestural scores for the English verbs *mistimes* and *mistakes* (based on acoustic data from Smith et al., 2012)

This provides an illustration of the gestural organization of tightness as well as a synchronic parallel for the functional relevance with which an historically more 'separate' morpheme becomes tighter to (in this case) the verb (possibly entrenching one collocation at a time). A similar trajectory of tightening may be reconstructed for the diachronic prefixation of the PIE spatial particles in Ancient Greek and other IE branches – thus providing a mechanism for the gradual split of the verbal prefix construction from that of the independent particle in terms of its tightness to the verb.

4. The diachrony of the spatial particle constructions in Ancient Greek

4.1 Spatial particles in Proto-Indo-European

This section sketches a scenario for the reconstruction and long-term construction developments of the spatial particles from PIE through Ancient Greek. The label *spatial particle* has been adopted here for the etyma of these morphemes in PIE and as a cover-term for the descendant morphemes in the constructions that develop in the different branches. Traditionally, the term 'preverb' has been applied to the PIE category as well as variously to some or all of the descendent morphemes, usually with some implicit functional claim to cross-linguistic 'category-hood'.[13]

From what is basically one construction in PIE, the spatial particles undergo a construction split into three distinct constructions by the time of Classical Greek: adposition, verbal prefix, and independent adverb. These are manifestly related but show functional distinctions and can, in Classical Greek, be combined with one another (cf. 3.1). This independence seems to develop through the course of the attested Greek material (cf. 4.2). Particularly notable is the development of abstract grammatical ('aspectual') meanings and lexicalizations in the prefix construction (see 4.3).

The spatial particles can be reconstructed for PIE as a relatively coherent group in the function of free adverbials and in some more or less close collocations with nouns and with verbs. Many of the reconstructed PIE spatial particles show clear traces of petrified nominal morphology: case endings, ablaut patterns, and derivational morphology.[14] Croft (2001, pp. 264f) describes the cross-linguis-

---

13. As for the conceptual confusion of essentialist linguistics (cf. Section 1, Croft, 2000, p. 24, passim), the 'preverb' label is a case in point: coined as a clipped shorthand to separate it from other 'prepositions', the term has since undergone reification (on which cf. Croft, 2000, p. 4, 2001, pp. 50f; Ohala, 1990, pp. 161f) into an *a priori* universal category (as evidenced by volumes dedicated to it, i.a. Rousseau, 1995; Booij & van Kemenade, 2003). This was opportune because the morphemes do not fit neatly as either derivational affixes or independent adverbials. Such nomenclature is arbitrary and even haphazard, yet it forms the conceptual basis for postulated universal categories. The terms can be traced to traditional, atheoretical, language-specific labels that may have been descriptive at some point for some data (ultimately from the Latin grammatical tradition, as also noted by Haspelmath, 2010). They have since been transposed – opportunistically (cf. Croft, 2001, pp. 30ff; Ohala, 1990, pp. 159ff) – as templates for other languages, accumulating a train of sundry meanings. Along the way, they have morphed into presumed theoretical–universal concepts – by the assumption of ontology from terminology, what Lakoff (1987, pp. 122f) calls the 'folk theory of reference'.

14. At least one has a clear etymology as a relational body-part noun: **h²ént-i* 'in front of', corresponding to Hittite *ḫant-s* 'forehead' (cf. Melchert, 2009, p. 613).

tically commonplace development of adpositions from relational nouns in terms of profile shift (cf. also Givón, 1989, pp. 54ff; Heine et al., 1991, passim; Bybee, 2010, pp. 138ff), as in e.g. English *on top of*. The situation in Ancient Greek (and other IE branches) stands out cross-linguistically Because the outcome of an etymon is usually both adposition and verbal prefix (and independent adverb) – and neither adposition nor prefix should be reconstructed as primary. An 'original' metaphorical development for these morphemes from a concrete noun to spatial and hence more abstract meanings would have taken place already in the earliest stratum of PIE: the reconstruction of the grammaticalized status of these morphemes as particles is certain (cf. i.a. Dunkel, 2014). The further development of the PIE spatial particle construction involves primarily a single syntagmatic element, and so is a split without prior chunking of the type of *gonna* (cf. 3.1).

A complication for the cross-linguistic comparison of a pathway from relational noun, which seems to assume juxtaposition, is that these morphemes are not specifically reconstructible as adpositions, but basically independent as to position. At any rate, *pace* Watkins (1963), postverbal position can also be reconstructed, cf. (16–17).

(16) Homeric Greek, Il. 2.699
ékhen káta gaĩa mélaina
hold.IMPF.3SG *káta* earth.NOM black.NOM
'the black earth held [him] *down*'

(17) Vedic Sanskrit, RV 10.105.10c
yáyā své pā́tre siñcása út
which.INSTR own.LOC cup.LOC pour.2SG.PRES.IND.MID *út*
'[the ladle] with which you pour *up* [drink] into your own cup'

The question is whether semantic changes emerged only in juxtaposition with other nouns or also independently and/or in conjunction with verbs. Arguably, the extensive variation in position and function of the particles should be reconstructed as primary, which agrees with the more general PIE tendency towards a high degree of word-order flexibility (cf. 2.2). This does not preclude the reconstruction of some specific word-order constructions, such as the frequent initial position (cf. MacDonell, 1916, p. 285 on Vedic). Anyhow, position does influence meaning. This again is the basic, iconic pattern that will develop into the prefix–preposition contrast: the closer the meaning is to affecting the noun, the closer it is placed relative to the noun, and vice versa (cf. also Givón, 2001, pp. 34f, 271ff). Such pragmatic iconicity may also include the relative importance of postposing and preposing and of utterance initial position etc. Up until Classical Greek, this is only a tendency: Homeric Greek includes patterns of non-adjacency even in the case of heavily lexicalized collocations.

Regardless of whether the particles developed originally from relational expressions, in the first instance they presumably served to specify events or referents in terms of location and direction. As word-order entrenched in the various frequent collocations that associated particles closely with position near noun or verb respectively, the specialized senses of these specific (iconic) uses emerged as individual constructions. Juxtaposition gradually petrified towards more and more semantic and formal 'tightness' (cf. 3.3).

Some original semantic contrast between the nominal and verbal associations – each could stem independently from the respective original metaphorical development from a concrete noun – would explain (some of) the huge variation of meaning to be reconstructed for many of the spatial particles.[15] As regards preposition contra postposition, MacDonell (1916, p. 285) notes that a functional distinction still exists in Vedic, comparable to the cross-linguistic tendency (Givón, 2001, p. 35) of emphasis in a fronted element. The consistent Classical Greek prepositioning is thus an inflationary development of 'bleaching' the construction of its pragmatic salience to the point where it becomes obligatory.

4.2 Construction developments in Ancient Greek

Like in many other IE branches, the outcomes of the PIE etyma in Classical Greek have crystallized into constructions of preposition (and postposition), verbal prefix and independent adverb. The formation of tight collocations with verbs and nouns respectively may have taken place one collocation at a time, a few of which could possibly be reconstructed to PIE. Seen in terms of construction split, out of a continuous span of token variation for the PIE construction, distinct clusters have emerged through the accumulation of continuous reconstrual of collocational patterns. The association of context–form with pragmatic interpretation has gradually entrenched so as to form distinct units.

The older attested languages – Homeric Greek, Vedic Sanskrit and Hittite – show ongoing developments independently in each branch. In the later stages of these languages (and in other later attested branches) this has advanced further – partly in different directions, partly in homologous developments. The ostensible teleology can be put to rest through a view of construction split informed by

15. Despite a traditional unwillingness to do so. One may note the various categorical proposals of rebracketing (cf. 2.1) that assume either version must be the original: Delbrück (1893, pp. 645ff) preverbal→postnominal, Schwyzer & Debrunner (1950, p. 430) postverbal→prenominal, Bubenik (2007) postnominal→preverbal. Again, cues are picked *ad hoc* for obligatory constituency, with the diachronic twist that the cues themselves first have to be reconstructed. Cf. Croft (2001, pp. 257f, 263).

Croft's (2000) *utterance selection* model (cf. 3.3): the process of grammaticalization, ultimately the movement of categorization spans in exemplar space, starts off as a pragmatic tendency, one which speaker selections are likely to continue to exploit, with each accumulating token increasing predictability and entrenchment. Thus the pre-stage of particle specialization is already present as a tendency in PIE, with particles generally positioned closer to elements to which they are perceived as pragmatically related. This gradually gains momentum separately in the branches, not because of a drive towards pre-established categories but because of the dialectic of form and function (cf. Haspelmath, 1998, pp. 321f, 345f), where the development in each generation is constrained by its initial state (cf. also Traugott & Trousdale, 2013, pp. 227ff).

The stage represented by Homeric Greek falls somewhere mid-way in the development from PIE to Classical Greek, with the substantial caveat that the status of the Homeric language is complex – it is a highly formulaic register and does not represent a single synchronic layer in the sense of the ordinary spoken language. The dating is highly unclear too: for convenience one may give a mid-7th century BCE date for the Iliad, following West (2011). Homeric Greek already exhibits distinguishable adposition and prefix constructions quite frequently: they exist side by side with the older pattern. It is unclear to what extent this is a product of the diachronic nature of the formulaic register; alternatively, it may bear resemblance to the state of affairs at some particular synchronic layer of the language outside the poets' variety.[16] Many of the tokens of the particles show considerable functional ambiguity (cf. i.a. Chantraine, 1953, pp. 82ff); this variety does not yet treat the particles exclusively in terms of three distinct units, though a substantial subset of instances are clear exponents of them. Frequent collocations will have grammaticalized further, with continuous input from the spoken language into the poetic register.

By the time of Classical Greek, the preposition and prefix constructions readily intersect with one another. This is seen in the 'double particle' construction, where one particle etymon appears simultaneously as both verbal prefix and as preposition (see Example 18b below), which is extremely unusual in Homeric Greek.[17] Arguably, this combination of two exponents of the same etymon only makes sense if prefix and preposition contribute different meanings, i.e. the expression represents the combination of two constructions (cf. 3.1), although

16. Horrocks' (1981) argument that this syntax must predate Mycenaean Greek (15th–12th centuries BCE) does not hold: the Mycenaean material is simply too limited for conclusions to be based on negative evidence.

17. Hessinger (1978, p. 218) adduces a couple of examples of Homeric prefix plus preposition, to which could be added prefix plus particle (e.g. Il. 23.709).

they may still have been perceived to some extent as being related.[18] A simple interpretation of the 'double particle' as an intensifying repetition of the particle does not account for the lexicalized meaning retained especially by the prefix; also, a simple doubling of either prefix or preposition does not occur. Further, preposition and prefix seem to alternate between different meanings and case constructions independently, just like different combinations of other prepositions and prefixes. There also seems to develop noncompositional functions specific to the 'double particle' (cf. 4.3).

The completed split of the constructions in Classical Greek relative to PIE is illustrated schematically in Figure 3, showing the constructions as 'budding off' and establishing themselves into niches (cf. Figure 1).

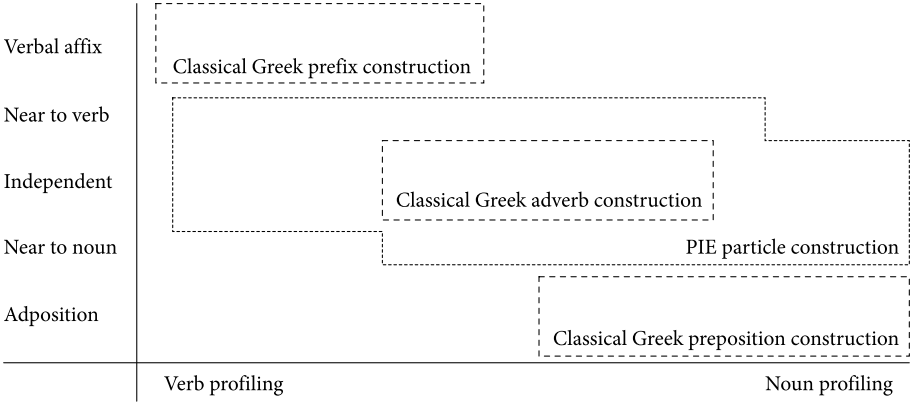


Figure 3. Schematic of the split of the spatial particle construction from PIE to Classical Greek

As an illustration of construction 'life-cycles' in a long-term perspective, the development of the spatial particle constructions shows several phases: the original metaphorical development(s) of adverbials from relational nouns; the clustering of tokens in different niches leading eventually to a split into separate constructions; the development of abstract and lexicalized senses; the eventual decay of the productivity of the non-lexicalized prefix construction, so that in Demotic Modern Greek, prefixes have become fully fused with the root (*phonogenesis*, Hopper, 1994), e.g. Classical Greek *ek=baínō* 'go out' to opaque Modern Greek *vyéno*.

18. Incidentally, Heine et al. (1991) present some experimental evidence on German prepositions vs. verbal particles, concluding that even the nominally 'identical' ones like *aus* are treated cognitively as quite distinct categories.

At one point in Ancient Greek it seems prefix and preposition constructions compete for expressing particular construals. The 'double particle' construction may possibly have originated in this competition. Another effect may be the increasing imposition of the functions of the constructions on the individual particles. This may be what is behind the form–function development seen in Tables 2–3, which summarizes data for major spatial sense divisions of the spatial particle *epi* in samples of Homeric and Classical Greek.

Table 2. Summary of Homeric Greek spatial senses of *epi* in Iliad 1, 15, 21

	Non-adposition		Adposition	
	Tokens	**Typical case**	**Tokens**	**Typical case**
Superposition	12	(acc, dat)	29	dat > acc > gen
Direction	28	dat	11	Dat
Proximity	8	(acc)	15	Dat
Attained direction	4	(dat)	10	Acc

Table 3. Summary of Classical Greek spatial senses of *epi* in Herodotus 7.1–99 (numbers in parentheses exclude tokens that were judged aspectual)

	Prefix		Preposition	
	Tokens	**Typical case**	**Tokens**	**Typical case**
Superposition	50 (24)	(acc, dat)	16	dat > gen
Direction	1 (1)	(acc)	48	Acc
Proximity	28 (7)	(acc, dat)	16	(dat, acc)
Attained direction	24 (12)	(acc)	12	Acc

Comparing Tables 2 and 3, these preliminary data suggest a major semantic shift of senses of *epi* across the constructions from Homeric to Classical Greek: from predominantly adposition for superposition 'on' and non-adposition for direction 'to', to predominantly preposition for direction 'to' and prefix for superposition 'on'.[19] This cannot be given full treatment here, but it may be suggested that it shows a trend towards increasing semantic weight residing in the constructions, in turn influencing the 'lexical' senses contributed by the particles. Such shifting of semantic locus is accountable in a construction grammar framework, as it sees constructions as semiotic units. This construction-driven semantic change correlates with the constructions becoming independent, whereby an increasing sense of contrast between them emerges. The contrast exhibited in

19. Additional data also indicate prefix connection with attainment in other senses than direction and with abstract senses.

Homeric Greek would be similar to that of PIE; in Classical Greek, the constructions have developed more independent semantics.

The development in *epi* could be traced to a PIE reconstruction where originally the sense of non-attained direction is where the particle is closest to the verb, and vice versa. In PIE, direction would be construed as corresponding more closely to the event than the goal, since the goal is not attained, and this situation is evidenced still by Homeric Greek. By Classical Greek, however, the novel construal focuses on the verbal event being attained with regard to the argument, as opposed to a more mediated 'direction towards' the argument. Compare Hopper & Thompson (1980, 1982) on semantic transitivity, especially the iconicity of how a preposition puts an argument "in a position more removed from the verb" (1982, pp. xiii–xiv). This change thus replaces one iconic motivation with another.

4.3 Grammaticalization of aspectuality in discourse

Not yet discussed are the aspectual functions of prefixation. Some such contrast does exist in Classical Greek and can be seen not least in the double particle construction, as in the alternation in (18a–b).

(18) a. Classical Greek, Hdt. 7.11
elásās *ep'* ándras toútous
march.AORIST.PTCP.NOM.SG *against* man.ACC.PL that.ACC.PL
'having marched against those men'

b. Classical Greek, Hdt. 7.9a
ep-elaúnōn *epì* toùs
against-march.PRES.PTCP.NOM.SG *against* the.ACC.PL
ándras toútous
man.ACC.PL that.ACC.PL
'marching upon those men'

One may briefly note that the prefix construction often comes with a completive sense (sometimes more akin to assertive realis) which could be connected to the semantic developments outlined in 4.2. Alternations of prefixed and non-prefixed verb in Greek in terms of aspectuality have received some discussion (i.a. Brunel, 1939; Watkins, 1967; Horrocks, 1981; Viti, 2008a, 2008b), but without consensus. The PIE spatial particles do turn up as the quintessentially perfective prefixes of Slavic (for one view, cf. Janda et al., 2013). Possibly, the Greek prefixes functionally resemble the preconditions for the Slavic development. Much the same goes for Sanskrit (Gonda, 1948), Germanic (e.g. Brinton, 1988), Latin (Delbrück, 1893). It may be suggested that a budding aspectuality can be reconstructed for

PIE, which then develops independently in the branches, correlated with the construction split (4.2).

One may note the growing recognition that many of the functions of 'aspectuality' are primarily discourse-related (Hopper, 1979a, 1979b; Fleischman, 1990; contributions in Hopper, 1982; and in Thelin, 1990), in particular in organizing narrative. Some of the nuance made in adding the prefix has been suggested in the translations for (18a–b), but such alternations should rather be viewed in their wider discourse context, which is beyond the scope here. It may be argued that it is often better to regard the Ancient Greek prefix as a *discourse construction* (cf. Imo, 2007; Lewis, 2011) which is added as a clitic to the verb in the relevant part of the discourse – straightforward in RCG because no fundamental distinction is made between discourse particle and derivational affix. This reflects the view (cf. Section 1) that the roots of grammaticalizations lie in creative discourse strategies, entrenching via less obligatory but still more or less conventionalized constructions that are used to express particular discourse effects (cf. i.a. Givón, 1979, p. 83, 2001, p. 36; Haspelmath, 1998, p. 346; Croft, 2000).

5. Conclusion

This paper has tried to show some ways in which RCG meets grammaticalization – in particular, on the fundamental issue of language change being a matter of degree. RCG has been presented as a way of looking at constructions (and hence language in general), in synchrony as well as diachrony, that is more in line with the nature of grammaticalization. Following the observation by Bybee (2010, p. 120), quoted in 3.1, on the need for a revised conception of synchronic grammar without covert units in order to capture the facts of grammatical change – namely that it is fundamentally a matter of degree – it has been argued here that RCG provides such a synchronic model which adequately addresses grammaticalization, and that it is also fruitful for modelling diachrony.

The diachronic development of the PIE spatial particles in Greek and other IE languages is conducive to this analysis. An RCG account opens for a more nuanced understanding of their diachrony, because it allows for gradient construction split and for treating their various formal cues as independent factors (especially, all syntactic developments are not seen in terms of linear order). There is no need, as the received wisdom would have it, to commit beforehand to one categorical model of reanalysis over another. Instead, construction split can be pursued as a pragmatics-driven gradual development involving reinterpretation, accumulation of discourse exemplars, and gradual activation of novel constructions, with no singular event of reanalysis of covert structures. With RCG, the

fundamental gradedness of synchrony and diachrony correspond to one another. The need for a comprehensive view of construction development has been emphasized: again, change stems from pragmatics, and it has been suggested that online phonology may provide a mechanism.

The completion of construction split can only be judged *post factum*. A relevant criterion is the stage where two now-distinct constructions can be combined with one another, implying two separate units. To the extent that the double particle construction is increasing from Homeric to Classical Greek, this seems indicative of increasing felicity of such combinations, thus a gradually developing distinctiveness between the nascent constructions. Further, not only do the functional distinctions between the constructions change, but the spatial senses of the particles are influenced by the evolving constructions. As the constructions grow apart, the meaning of the lexical items that instantiate them changes. The particular semantics of the double particle construction seems to suggest that it too is an independent construction by Classical Greek, derived from the pragmatic use of the prefix and preposition constructions together.

In exploring the theoretical groundwork for these developments, the article has advocated the RCG view that all linguistic structures are language-specific constructions, defined in terms of cues. It follows there is no language-independent structural analysis, and because such discrete analyses are not necessary in order to define utterances, there is no necessary discreteness in language change. Seen from the perspective of construction grammar, the concept of grammaticalization may require redefinition, since if any change is a change in constructions, what changes can be considered 'more grammatical' is open-ended. But it is as important to acknowledge, as RCG does, the basic finding of grammaticalization: that change is fundamentally gradual.

References

Bakker, E. J. (2010). Pragmatics: Speech and Text. In E. J. Bakker (Ed.), *A Companion to the Ancient Greek Language* (pp. 151–67). Chichester: Wiley-Blackwell. https://doi.org/10.1002/9781444317398.ch11

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89(1), 1–47. https://doi.org/10.1037/0033-295X.89.1.1

Booij, G., & van Kemenade, A. (Eds.). (2003). *Yearbook of Morphology 2003: Preverbs*. New York: Kluwer Academic. https://doi.org/10.1007/978-1-4020-1513-7

Bolinger, D. L. (1961). *Generality, Gradience, and the All-or-None*. 's-Gravenhage: Mouton.

Brinton, L. J. (1988). *The development of English aspectual systems: aspectualizers and post-verbal particles*. Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3–4), 155–180. https://doi.org/10.1159/000261913

Brunel, J. (1939). *L'aspect verbal et l'emploi des préverbes en grec, particulièrement en Attique*. Paris: Klincksieck.

Bubenik, V. (2007). On the evolution of the Greek prepositional system from Proto-Indo-European local adverbs and adpositions. In G. K. Giannakis et al. (Eds.), *Proceedings of the 8th International Conference of Greek Linguistics*. University of Ioannina.

Bybee, J., & Hopper, P. J. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45

Bybee, J. (2004). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Chantraine, P. (1953). *Grammaire homérique. Tome II: syntaxe*. Paris: Klincksieck.

Chatterjee, R. (1982). On Cross-Linguistic Categories and Related Problems. In P. J. Hopper (Ed.), *Tense-Aspect: Between Semantics & Pragmatics* (pp. 335–345). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.1.20cha

Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow: Longman.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Croft, W. (2005). Logical and typological arguments for Radical Construction Grammar. In J.-O. Östman, & M. Fried (Eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (pp. 273–314). Amsterdam: John Benjamins. https://doi.org/10.1075/cal.3.11cro

Croft, W. (manuscript) *Exemplar semantics*. Draft, September 2007. Retrieved on 9 October 2017 from www.unm.edu/~wcroft/Papers/CSDL8-paper.pdf

Croft, W. (2008). On iconicity of distance. *Cognitive Linguistics*, 19(1), 49–57. https://doi.org/10.1515/COG.2008.003

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803864

Cummins, F. (2011). Coordination, not control, is central to movement. In A. Esposito et al. (Eds.), *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues* (pp. 243–255). Berlin/Heidelberg: Springer. https://doi.org/10.1007/978-3-642-18184-9_20

Debrunner, E., & Schwyzer, A. (1950). *Griechische Grammatik. Bd. II: Syntax und syntaktische Stilistik*. München: Beck'sche Verlagsbuchhandlung.

Delbrück, B. (1893). *Vergleichende Syntax der indogermanischen Sprachen, Erster Theil*. Strassburg: Trübner.

Denniston, J. D. (1952). *Greek Prose Style*. Oxford: Clarendon.

Dik, H. (1995). *Word order in Ancient Greek: A pragmatic account of word order variation in Herodotus*. Amsterdam: Gieben.

Dover, K. J. (1960). *Greek word order*. Cambridge: Cambridge University Press.

Dunkel, G. E. (2014). *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Heidelberg: Winter.

Fenk-Oczlon, G. (2001). Familiarity, information flow, and linguistic form. In J. Bybee, & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 432–448). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45.22fen

Fischer, A. (1999). What, if Anything, is Phonological Iconicity? In M. Nänny, & O. Fischer (Eds.), *Form miming meaning* (pp. 125–134). Amsterdam: John Benjamins. https://doi.org/10.1075/ill.1.12fis

Fischer, O. (1999). On the Role Played by Iconicity in Grammaticalisation Processes. In M. Nänny, & O. Fischer (Eds.), *Form miming meaning* (pp. 251–284). Amsterdam: John Benjamins. https://doi.org/10.1075/ill.1.20fis

Fleischman, S. (1990). *Tense and Narrativity: From Medieval Performance to Modern Fiction*. London: Routledge.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Givón, T. (1988). The pragmatics of word order: Predictability, importance and attention. In M. Hammond, E. A. Moravcsik, & J. Wirth (Eds.), *Studies in Syntactic Typology* (pp. 243–284). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.17.18giv

Givón, T. (1989). *Mind, code and context: Essays in pragmatics*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Givón, T. (1991). Isomorphism in the grammatical code: cognitive and biological considerations. *Studies in language*, 15(1), 85–114. https://doi.org/10.1075/sl.15.1.04giv

Givón, T. (1995). Isomorphism in the grammatical code. In R. Simone (Ed.), *Iconicity in language* (pp. 47–76). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.110.07giv

Givón, T. (2001). *Syntax: an introduction*. Volume 1. Amsterdam: John Benjamins.

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

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Gonda, J. (1948). Quelques observations sur l'emploi de verbe simple 'au lieu d'un composé' etc. dans la langue sanskrite. *Acta Orientalia*, 20, 167–205.

Harris, A. (1981). *Georgian Syntax: A Study in Relational Grammar*. Cambridge: Cambridge University Press.

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in language*, 22(2), 315–351. https://doi.org/10.1075/sl.22.2.03has

Haspelmath, M. (2008a). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1), 1–33. https://doi.org/10.1515/COG.2008.001

Haspelmath, M. (2008b). Reply to Haiman and Croft. *Cognitive Linguistics*, 19(1), 59–66. https://doi.org/10.1515/COG.2008.004

Haspelmath, M. (2010). Framework-free grammatical theory. In B. Heine, & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 341–365). Oxford: Oxford University Press.

Heine, B., Claudi, U. , & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Hessinger, J. J. (1978). The syntactic and semantic status of prepositions in Greek. *Classical Philology*, 73(3), 211–223. https://doi.org/10.1086/366433

Hirsch, E. D. (1967). *Validity in interpretation*. New Haven: Yale University Press.

Holmes, J., & Hudson, R. (2005). Constructions in word grammar. In J.-O. Östman, & M. Fried (Eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (pp. 243–272). Amsterdam: John Benjamins. https://doi.org/10.1075/cal.3.10hol

Hopper, P. J. (1979a). Some observations on the typology of focus and aspect in narrative language. *Studies in Language Groningen*, 3(1), 37–64. https://doi.org/10.1075/sl.3.1.03hop

Hopper, P. J. (1979b). Aspect and Foregrounding in Discourse in Discourse and Syntax. In T. Givón (Ed.), *Syntax and Semantics, Volume 12: Discourse and syntax* (pp. 213–241). New York: Academic Press.

Hopper, P. J. (Ed.). (1982). *Tense-aspect: between semantics & pragmatics*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.1

Hopper, P. J. (1991). On some principles of grammaticization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to grammaticalization*, Volume 1 (pp. 17–35). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.19.1.04hop

Hopper, P. J. (1994). Phonogenesis. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization* (pp. 29–45). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.109.09hop

Hopper, P. J. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 155–75). Mahwah, N.J.: Erlbaum.

Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Horrocks, G. C. (1981). *Space and time in Homer. Prepositional and adverbial particles in the Greek epic*. Salem, NH: Ayer Company.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor Adaptation of Speech I: Compensation and Adaptation. *Journal of Speech, Language, and Hearing Research*, 45(2), 295–310. https://doi.org/10.1044/1092-4388(2002/023)

Imo, W. (2007). Der Zwang zur Kategorienbildung: Probleme der Anwendung der Construction Grammar bei der Analyse gesprochener Sprache. *Gesprächsforschung–Online-Zeitschrift zur verbalen Interaktion*, 8, 22–45.

Janda, L., Endresen, A., Kuznetsova, J., Lyashevskaya, O., Makarova, A., Nesset, T., & Sokolova, S. (2013). *Why Russian aspectual prefixes aren't empty: prefixes as verb classifiers*. Bloomington, Indiana: Slavica Publishers.

Johnson, K. (2001). Spoken language variability: Implications for modeling speech perception. In R. Smits, J. Kingston, T. M. Nearey, & R. Zondervan (Eds.), *Proceedings of the Workshop on Speech Recognition as Pattern Classification (SPRAAC)*. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved on 9 October 9 2017 from http://linguistics.berkeley.edu/~kjohnson/papers/JohnsonSpraac.pdf

Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. In S. Maria-Josep, P. Speeter Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology* (pp. 25–40). Oxford: Oxford University Press.

Krug, M. G. (2001). Frequency, iconicity, categorization: Evidence from emerging modals. In J. Bybee, & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 309–335). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45.16kru

Kuningas, J., & Leino, J. (2006). Word orders and construction grammar. In M. Suominen et al. (Eds.), *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday* (pp. 301–309). Turku: Linguistic Association of Finland.

Ladefoged, P. (1990). Some reflections on the IPA. *UCLA Working Papers in Phonetics*, 74, 61–76.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago. https://doi.org/10.7208/chicago/9780226471013.001.0001

Lewis, D. M. (2011). A discourse-constructional approach to the emergence of discourse markers in English. *Linguistics*, 49(2), 415–443. https://doi.org/10.1515/ling.2011.013

Lichtenberk, F. (1991). Semantic change and heterosemy in grammaticalization. *Language*, 67(3), 475–509. https://doi.org/10.1353/lan.1991.0009

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-009-2037-8_16

Lindblom, B., Guion, S., Hura, S., Moon, S. J., & Willerman, R. (1995). Is sound change adaptive? *Rivista di linguistica*, 7, 5–36.

Lorenz, D. (2012). On-going change in English modality: Emancipation through frequency. *Zeitschrift für Literaturwissenschaft und Linguistik*, 169(43), 33–48.

Lorenz, D. (2013). From reduction to emancipation: is *gonna* a word? In H. Hasselgård, J. Ebeling, & S. Oksefjell Ebeling (Eds.), *Corpus Perspectives on Patterns of Lexis* (pp. 133–152). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.57.11lor

Macdonell, A. A. (1916). *A Vedic grammar for students*. Oxford: Oxford University Press.

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, 14(3), 380–407.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, 25, 71–102. https://doi.org/10.1016/0010-0277(87)90005-9

Matić, D. (2003). Topic, focus, and discourse structure: Ancient Greek word order. *Studies in Language*, 27(3), 573–633. https://doi.org/10.1075/sl.27.3.05mat

Melchert, H. C. (2009). Local Adverbs in Hittite: Synchrony and Diachrony. *Language and Linguistics Compass*, 3(2), 607–620. https://doi.org/10.1111/j.1749-818X.2009.00132.x

Morton, K. (1992). Pragmatic phonetics. *Advances in Speech, Hearing and Language Processing*, 2, 17–53.

Ohala, J. J. (1990). There is no interface between phonology and phonetics: a personal view. *Journal of phonetics*, 18(2), 153–172.

Rousseau, A. (1995). *Les préverbes dans les langues d'Europe: introduction à l'étude de la préverbation*. Villeneuve d'Ascq: Presses universitaires du Septentrion.

Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, 40(5), 689–705. https://doi.org/10.1016/j.wocn.2012.04.002

Tatham, M. (1990). Cognitive phonetics. *Advances in Speech, Hearing and Language Processing*, 1, 193–218.

Tatham, M, & Morton, K. (2006). *Speech production and perception*. New York: Palgrave Macmillan. https://doi.org/10.1057/9780230513969

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Thelin, N. B. (Ed.). (1990). *Verbal aspect in discourse*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.5

Vamling, K. (1989). *Complementation in Georgian*. Lund: Lund University Press.
Vihman, M., & Croft, W. (2007). Phonological development: Toward a "radical" templatic phonology. *Linguistics*, 45(4), 683–725. https://doi.org/10.1515/LING.2007.021
Viti, C. (2008a). From space words to transitive markers: the case of ancient Greek *en* 'in'. *Transactions of the Philological Society*, 106, 375–413. https://doi.org/10.1111/j.1467-968X.2008.00204.x
Viti, C. (2008b). Coding spatial relations in Homeric Greek: preverbs vs. prepositions. *Historische Sprachforschung*, 121, 114–161.
Watkins, C. (1963). Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb. *Celtica*, 6, 1–49.
Watkins, C. (1967). An Indo-European Construction in Greek and Latin. *Harvard Studies in Classical Philology*, 71, 115–119. https://doi.org/10.2307/310758
West, M.L. (2011). *The making of the Iliad*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199590070.001.0001

Construction index

Dutch
krijgen passive 33
partitive genitive construction 34

English
binominal construction 138, 141–142
caused-motion construction 110
copular construction 45, 51, 56, 58–59, 61, 63, 65, 70
degree modifier construction 141–142,
discourse marker *like* 30
ditransitive construction 23, 30
get passive 27, 30
going to future 5, 7, 141, 290
let alone construction 23
locational construction 45, 51, 56–58
locative construction 44
parenthetical concessive clause 24
partitive construction 137,
postural construction 45, 55
that clause 27–28
to infinitive 30
[*what's* X*doing* Y] construction 31

Georgian
ergative case construction 286

German
degree modifier construction 9, 15, 139, 143–145, 152–153, 156–162
partitive construction 144
quantifier construction 160

Greek
double particle construction 296, 301–304, 306
nominal agreement construction 283
prefix construction 301, 304–306
preposition construction 296, 301, 303, 306

Latin
[*secundum* NP] construction 7, 15, 169–171, 173–176, 178, 180–187, 189–193, 197–198

Manchu, Mongolian
attributive construction 252, 264
  attributive privative construction 271
auxiliary construction 257
comparative construction 258, 271
locational construction 252
negative construction 270–271
  negative copula construction 271
  negative existential construction 16, 248, 255, 259, 270
  negative obligation construction 266, 271
  prohibitive construction 255
  standard negation construction 16, 242, 251, 253, 258, 271

Norwegian
coordination construction 6, 9, 81, 89, 92
pseudocoordination 14, 75–91, 93, 98–99, 102–103
serial verb construction 85
verb phrase coordination 9, 75–76, 81, 103

Spanish
binominal construction 160
caused-motion construction 111
complex conjunction construction 206, 213, 215, 231, 235
conditional *caso* construction 15–16, 205, 207–210, 213–214, 216–217, 220–222, 227–235
locative construction 119
inchoative construction 6–7, 9, 14, 108, 110–130
nominal complement clause construction 224–227, 230–231, 235

Proto-Indo-European
spatial particle construction 16, 279, 298–299, 302

Index

A

activation 31–33, 225, 279, 281, 283, 291–291, 295, 305
adjacency 156, 282, 283, 287, 290, 299
adposition 16, 180, 295–296, 298–299, 301, 303
Afrikaans 76
agreement 78, 155, 282–287
Alchuka 246–247, 249–251, 256, 257
allostructions 114, 120, 207, 233–235
Amuric 267
analogy 10, 122, 139, 141, 159, 161–162, 172, 174, 180, 189–190, 192–193, 243, 289–290
Ancient Greek 16, 278–279, 282–285, 287–290, 296–300, 303, 305
areal construction grammar 241, 269–270
areal linguistics 241, 245, 270
aspect 44, 46, 51–53, 56, 77, 83–84, 86, 89, 98, 103, 107, 109, 111, 118, 126, 257, 269, 298, 303–304
asymmetry 81, 257, 263, 284, 296

B

Bala 246–247, 250–251, 260
Basque 285
bleaching 4, 14, 45–47, 65, 77, 85–87, 89–91, 93, 98, 103, 108, 137, 140, 293, 300
Bonan 261–262
bonding 123, 256
bracketing 160, 283, 287, 300
bridging context 10, 108, 139, 154, 157, 210, 218, 220
Buryat 261–264

C

categorization 139, 144, 283, 287, 291, 295, 301
caused-motion 110–111, 120
Chinese Pidgin Russian 267
chunking 146, 148, 281, 290, 292, 295, 299
clitic 12, 24, 108, 264, 287–288, 292, 295–296, 305
coherence 179–180, 192
collocation 6, 9, 33–34, 60, 64, 70, 177, 191–193, 195, 264–265, 284, 295–297, 300–301
collostructional analysis 77, 91
collostructional strength 93, 95, 99
compositionality 6, 26, 115, 125, 130, 178–179, 210, 221, 227, 235
compound prepositions 206, 208, 211–214, 231, 296
connection 13, 26, 30–35
constructional change 3, 12–13, 15–16, 22, 24, 26–29, 31, 35, 112, 117, 121, 139, 162, 170, 183, 187, 197–198, 209–210, 212, 229, 233, 242, 248, 254, 270
constructional network 8, 13, 16, 21–22, 26, 29–31, 33–35, 110, 140, 157, 190–191, 197, 230–231, 234–235, 241–242, 245, 251, 259, 270
constructionalization 6, 12, 15–16, 22, 26–29, 31, 35, 112, 140, 205, 207–210, 212, 216, 221, 229–230, 233, 235, 241–242, 244–245, 248, 252, 258, 269–270
contentful 27, 78, 140, 242, 270, 293
context expansion 7, 139–140, 160, 162
coordinate Structure Constraint 77
coordination 6, 9, 76–77, 79, 81, 89–90, 92, 103, 293
copula 7, 13, 43–47, 50–52, 55–56, 58–61, 63–65, 68–71, 80, 115, 120, 127, 248, 252, 254, 257, 265, 266, 268, 271

D

Dagur 261–262, 264, 271
Danish 76, 83
dative 23–24, 252, 266, 284, 286
decategorialization 75, 77, 88, 155, 160, 210, 221, 229
desemanticization 210, 220–221, 229
diachronic construction grammar 4, 6, 10, 13, 17, 21–26, 33–36, 137, 139, 209, 241–242
discourse 5, 12, 16, 30, 115, 141, 157, 160, 277, 279, 281–284, 288, 290–293, 296–297, 304–305
Dunan 267, 271
durative 44, 84, 126
Dzungaria 246

E

emergence 21–22, 26–31, 34–35, 116, 137, 160, 173, 177, 248, 257–258, 290
English 7, 13, 15, 24, 26–28, 30, 34–35, 43–56, 58–60, 62, 64–65, 67–71, 76–77, 81–83, 86, 90, 93, 97, 137–139, 141–142, 144–145, 152, 157, 160–162, 187, 206, 218, 234, 278, 283, 287, 296–297, 299
entrenchment 6, 22, 33, 46, 52, 58, 71, 162–163, 170, 177, 297
Even 246, 253–254, 256
Evenki 243, 246, 248, 260, 270
Ewenic 246, 248–249,253

exemplar theory 59, 278, 291

F
Faroese 76
Fisher's exact test 93, 121
frequency 7, 15, 26, 30, 34, 46, 56–58, 63, 68, 87, 93, 95, 97, 99, 112, 114, 117, 119, 123, 125, 130, 137, 139, 141, 145, 148–149, 160, 163, 169–171, 177, 198, 223, 234, 256, 292–293
frequency effects 10, 177
frequency, token see *token frequency*
frequency, type see *type frequency*
frequency, relative see *relative frequency*

G
genitive 34, 287
Georgian 285–287
Germanic 13, 44–45, 47, 50, 52–54, 76, 91, 304
grammaticalised islands 43, 46, 70

H
hapax (legomena) 7, 58–59, 93, 171, 178–183
head 24, 142, 155, 183, 188, 224, 282, 287, 295
Hittite 298, 300
host-class expansion 5–9, 15, 46, 59, 169–171, 177–178, 183, 187, 191–192, 197, 229

I
iconicity 159, 279, 288, 293, 299, 304
imperative 78–79, 92
imperfective 84, 248–250, 253, 255–257, 261
infinitive 7, 30, 34, 55, 78, 92, 99, 108–111, 115, 119, 121, 123–124, 126–128, 284
inheritance 31–32, 34, 110, 120, 138, 231, 242, 288

J
Japonic 267
Jurchen 246–247, 249–251, 254, 262
Jurchenic 246–249, 251, 257, 262, 264, 268, 271

K
Kazakh 245, 268
Khalkha 241, 261–262, 264
Khitan 261
Khitano-Mongolic 261
Khorchin 261–262
Kilen 246, 249–251
Latin 7, 9, 15, 55, 114, 118, 122, 169, 171–175, 178, 180–186, 188–193, 195, 197, 206, 211, 214–215, 249, 295, 298, 304

L
layering 156, 160
location 44–52, 55–58, 61, 63, 71, 80, 87–88, 108, 111
locative 13, 43–47, 50–54, 56–57, 62, 65, 68–70, 83, 88, 102–103, 108, 110–111, 115–116, 119, 121–123, 126, 130, 206, 217, 252
Low German 76

M
macro-construction 141
Manchu 16, 241, 243–260, 262–266, 269–272
Mandarin 245, 249, 266, 269
Mangghuer 261–262
meso-construction 141, 152–153, 157, 159, 161–162
micro-construction 8–9, 15, 27, 71, 109–110, 114, 123, 141–142, 156–157, 159–163, 214, 227, 230–231, 233–235, 258
Mongolian 16, 241, 243–245, 248, 261, 263–264, 266, 270–272
motion 7, 44, 51–53, 61, 63, 83, 87–88, 92–93, 95, 97, 99, 102–103, 260
motion, atelic 14, 75, 77, 80, 86–87, 90–91, 98, 102–103
motion, body-part 80, 93
motion, telic 80, 82, 84, 86, 98–99

N
Nanaic 246–249, 251
negation 16, 143, 241–243, 248–249, 251, 253, 257–259, 261, 263–264, 266, 268–271
Negidal 246, 260
Nivkh 267–268
node changes 30–31
nodes 13, 21–22, 26, 31–33, 35, 80, 110, 157, 209, 231, 233, 258, 295
nominative 51, 283, 286
Norwegian 6, 9, 14, 75–78, 80, 84–86, 88, 90, 98, 102
noun phrase 28, 155, 252

O
Oirat 260–262
Old Georgian 287
Ordos 261–262
Oroqen 249–250

P
paradigm 7–10, 25, 79, 84–85, 87, 89, 91, 100, 102, 110–111, 138, 152, 156–157, 163, 171–172, 211, 214–215, 221, 227, 230, 279, 283
paradigmatization 75, 77, 87, 89, 102, 137, 160
participle 33, 172, 212, 248–249, 254–257, 259, 262–263, 265, 268
participle, perfect 78, 99, 253, 255–256
participle, present 78, 92, 172
phrase structure 281, 285, 289
polyfunctionality 140, 156
polysemy 140
possession 248, 252
postural 45–47, 55–56, 61–62, 65, 67–69, 71
posture verbs 7, 12–14, 43–45, 47, 52–54, 71, 84–85, 87
pragmatics 16, 142, 222, 277, 279, 293, 305–306
predictability 293–294, 301

procedural 10, 16, 27, 78, 89, 108, 173, 242, 270
productivity 7, 13–15, 25–26, 31, 46, 58–61, 93, 117, 126, 130, 160, 162, 169–171, 174, 177–183, 192–193, 197, 210, 221, 235, 302
progressive 43–44, 53, 56, 84
Proto-Indo-European 16, 277, 298
prototype 7, 9, 128, 130, 160, 191, 279, 291, 294–295
pseudocoordination 14, 75–91, 93, 98–99, 102–103

Q
Qapqal 257

R
re-wiring changes 31
reanalysis 16, 108–109, 116, 137, 142, 154–155, 160, 173, 211, 253–254, 289–290, 305
reconstruction 108, 171, 260, 298–299, 304
reduction 6, 9–10, 15, 32, 87, 108, 139–140, 154, 162, 170, 182, 268, 290, 292–295, 297
relabeling 260, 270
relative frequency 7, 112, 117–119, 121, 124, 152, 154, 170, 261
relative pronoun 30, 284

S
Sanskrit 299–300, 304
Scandinavian languages 44, 75–77
schematicity 7–10, 15, 25–26, 92, 110, 116, 129, 140, 171, 177–178, 197, 210, 221, 227, 235, 241–242, 252, 263
schematization 254
Selkup 268
single-event meaning 14, 82–83, 85–86, 88, 90, 103
Solon 246, 249–250
Spanish 6, 14–15, 107, 109, 111–115, 117–119, 122, 124–125, 128–130, 138, 160, 205–209, 211–217, 224, 227, 230, 233–235
strengthening 6, 30–31, 90, 140, 163
subevent 82
subordinating 211–214, 227
subschema 14, 76–78, 80, 84–85, 87, 90–94, 96–102, 110–111, 123, 138, 157, 161–162, 231
supine 78
Swedish 44–45, 47, 50, 54, 76, 78, 80, 84, 86–87, 206, 230
switch context 108
syntactic incorporation 6, 115, 123–124, 130
syntactic expansion 177–178, 183, 187, 192

T
tightness 282, 291–292, 296–297, 300
token frequency 7, 112, 117–118, 147, 149, 163, 177–179, 183, 190–191, 261
Tungusic 242, 244–248, 250–251, 255–258, 260–262, 267–268, 271
Turkic 268
type frequency 7, 15, 93, 170, 177–180, 183, 192–193

U
Udegheic 246, 248
Udihe 243, 246, 248, 260
Uilta 246–248, 251, 256,

V
Vedic 299–300
verb phrase 9, 75–76, 78, 81, 103, 286

W
word order 5, 112, 114–115, 123, 269, 272, 280, 285, 288–289
Wutun 269

Y
Yakut 245, 268

Grammaticalization research has increasingly highlighted the notion of constructions in the last decade. In the wake of this heightened interest, efforts have been made in grammaticalization research to more precisely articulate the largely pretheoretical notion of construction in the theoretical framework of construction grammar. As such, grammaticalization research increasingly interacts and converges with the emerging field of diachronic construction grammar. This volume brings together articles that are situated at the intersection of grammaticalization research and diachronic construction grammar. All articles share an interest in integrating insights from grammaticalization research and construction grammar in order to advance our understanding of empirical cases of grammaticalization. Constructions at various levels of abstractness are investigated, both in well-documented languages, such as Ancient Greek, Latin, Spanish, German, Norwegian and English, and in less-described languages, such as Manchu and Mongolian.

ISBN 978 90 272 0064 8

John Benjamins Publishing Company