

DE GRUYTER
MOUTON

Robert Mailhammer (Ed.)

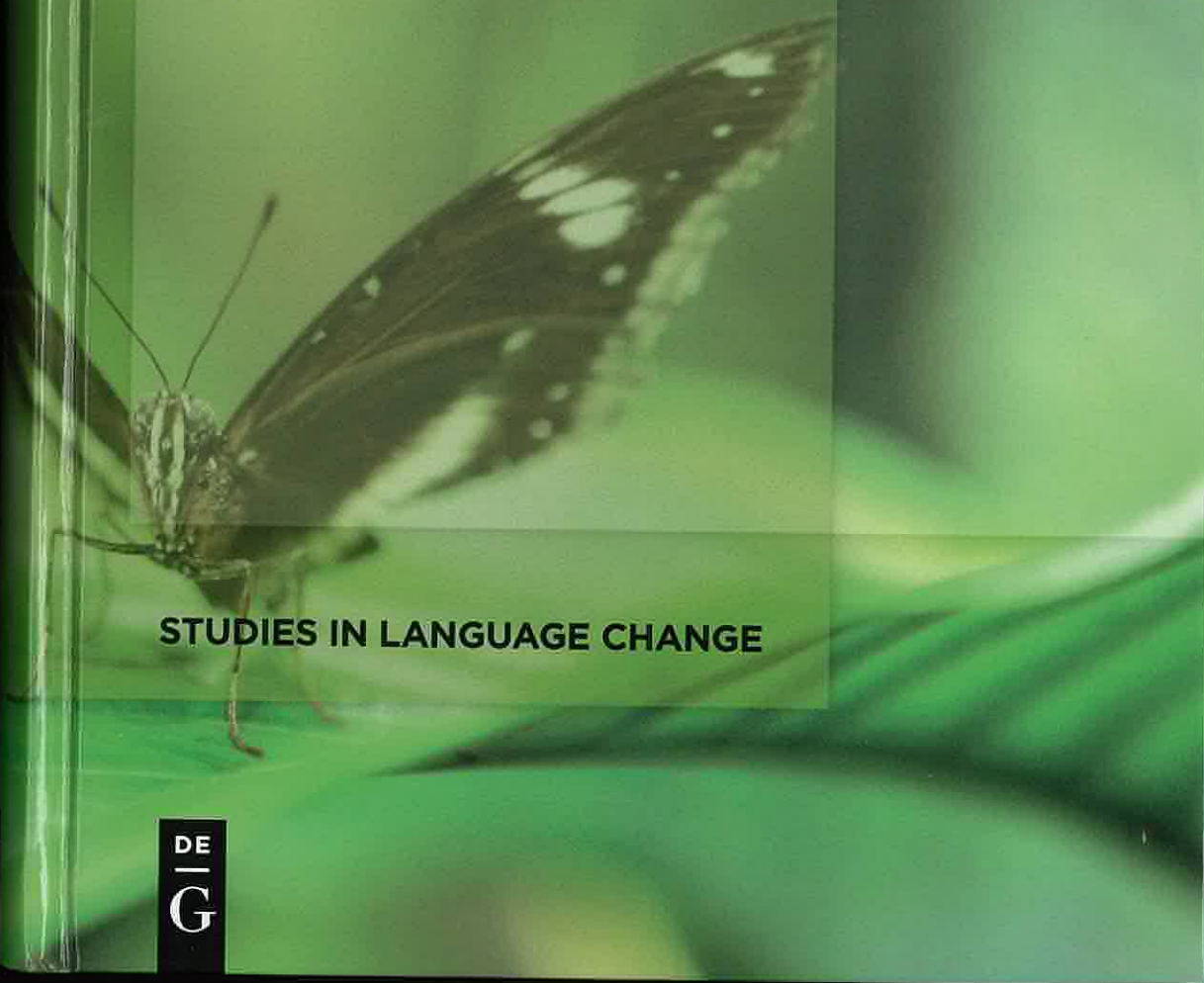
LEXICAL AND STRUCTURAL ETYMOLOGY

BEYOND WORD HISTORIES

STUDIES IN LANGUAGE CHANGE

DE
|
G

Robert Mailhammer (Ed.)



**Studies in
Language Change**

Edited by
Cynthia Allen
Harold Koch
Malcolm Ross

Volume 11

Beyond Word Histories

**Lexical
and Structural
Etymology**

Beyond Word Histories

Edited by
Robert Mailhammer

**DE GRUYTER
MOUTON**

Studies in
Language Change
and
Etymology

Edited by
Judith Allen
David Nash
Patrick Doss

Volume 11

ISBN 978-1-61451-059-8
e-ISBN 978-1-61451-058-1
ISSN 2163-0992

Library of Congress Cataloging-in-Publication Data

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

Bibliographic information published by the Deutsche Nationalbibliothek

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

© 2013 Walter de Gruyter, Inc., Boston/Berlin
Cover image: iStockphoto/thinkstock
Typesetting: RoyalStandard, Hong Kong
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen
♻️ Printed on acid-free paper
Printed in Germany

www.degruyter.com



Contact details and affiliations of contributors

Robert Mailhammer, University of Western Sydney
R.Mailhammer@uws.edu.au

Luise Hercus, Australian National University
luiseh@dodo.com.au

Harold Koch, Australian National University
Harold.Koch@anu.edu.au

Meredith Osmond, Australian National University
Meredith.Osmond@anu.edu.au

Patrick McConvell, Australian National University
Patrick.McConvell@anu.edu.au

Theo Vennemann, Ludwig-Maximilians Universität München
Vennemann.Theo@t-online.de

Alexandre François, CNRS–LACITO; Labex 'Empirical Foundations of Linguistics' Australian National University
francois@vjf.cnrs.fr

Rachel Hendery, Australian National University
Rachel.Hendery@anu.edu.au

David Nash, Australian National University
David.Nash@anu.edu.au

John Giaccon, Australian National University
jgiacon@ozemail.com.au

WM	wife's mother		
Wm	Wangkumara		
WN	Wangaaybuwan		
WTS	Western Torres Strait		
Ya	Yandhrundwandha		
YG	Yuwaalaraay Gamilaraay		
Yi	Yarluyandi		
YR	Yuwaalaraay (YR)		
Yw	Yawarrawarka		
YY	Yuwaalayaay		

Table of contents

Contact details and affiliations of contributors — v
 Abbreviations — vii

Robert Mailhammer
 Introduction: Etymology beyond word histories — 1

Robert Mailhammer
 Towards a framework of contact etymology — 9

Harold Koch and Luise Hercus
 Obscure vs. transparent cognates in linguistic reconstruction — 33

Harold Koch
 The etymology of a paradigm: the Pama-Nyungan 3SGF reconsidered — 53

Rachel Hendery
 Constructional etymology: The sources of relative clauses — 83

Theo Vennemann
 Concerning *myself* — 121

Patrick McConvell
 Granny got cross: semantic change of *kami* 'mother's mother' to 'father's mother' in Pama-Nyungan — 147

Alexandre François
 Shadows of bygone lives: The histories of spiritual words in northern Vanuatu — 185

Meredith Osmond
 Markers of the spirit world in Oceanic languages — 245

John Giacon
 Etymology of Yuwaalaraay Gamilaraay bird names — 251

David Nash
 The smuggled budgie: case study of an Australian loanblend — 293

Luise Hercus
 Archaisms in placenames in Arabana-Wangkangurru country — 313

Subject index — 323

Robert Mailhammer

Introduction: Etymology beyond word histories

1 Etymology and historical linguistics

Etymological research naturally is fundamental to the historical investigation of a language. Etymologies permit generalizations about historical developments, for instance the formulation of sound laws, or about a particular synchronically attested phenomenon, such as the occurrence of a particular type of stem formation. Moreover, etymologies are also used as arguments to answer questions about historical relationships among languages.

Consequently, the historical investigation of a language generally begins with a comparative analysis of single lexemes, which leads to etymologies via the reconstruction of a protoform. Etymologies then serve as basis for systematic generalizations. Conversely, etymological data are also used to verify hypothesized historical connections. In these cases etymologies do not serve a self-fulfilling purpose, i.e. to find out about the origin of a word, rather they are arguments in a historical discussion. As a result, etymology can be regarded as something like a fundamental auxiliary discipline of historical linguistics. However, etymological research can be focused more narrowly on aspects of historical word formation. From this perspective, more general statements about historical developments in a language are not explicitly part of the investigation. Rather, the aim is to pinpoint the origin of a word, its *Urschöpfung*, as Seebold (1980: 431) points out.

Wenn wir ein Wort etymologisch untersuchen, dann richten wir unser Hauptaugenmerk auf seine Entstehung: Wir suchen zu zeigen, daß es aus einem nachweisbaren Grundwort nach dem Muster eines ebenfalls nachweisbaren Wortbildungstyps geprägt wurde. Solange der Zusammenhang zwischen Grundwort und abhängigem Wort noch voll durchsichtig ist, gehört diese Fragestellung in die Wortbildungslehre der betreffenden Sprachstufe; sobald aber geschichtliche Belege und geschichtliche Überlegungen eine Rolle spielen, befinden wir uns im Bereich der Etymologie. (Seebold 1980: 431)

[If we examine a word etymologically, we focus on its origin. We attempt to show that it was coined based on an evidenced source lexeme following the model of a type of word formation that is likewise attested. As long as the connection between the base word and the derived word is still fully transparent, this is a problem belonging to the domain of word formation of the language examined; as soon as historical attestations and historical considerations come to play a role, we are in the domain of etymology.]¹

¹ All translations are mine unless indicated otherwise.

Usually, such studies are undertaken once sufficient historical knowledge has been acquired, as the investigation of a word's history is based on generalizations and observed developments. However, detailed studies of a word's history are subsequently used to test hypothesized generalizations, which in turn leads to improved etymologies and so forth.

As a result, there are in principle two applications of etymological research, which differ in their immediate aims. One is focused on the linguistic history of a language and the other on the linguistic history of a linguistic element. In practice, however, these two perspectives cannot always be strictly separated, because every etymology naturally points beyond the mere origin of a word. But conceptually it makes sense to differentiate between the two, without claiming that this separation is necessary rigid in practice.

2 The etymological approach

The separation between etymology as an auxiliary discipline of historical linguistics and a discipline in its own right focusing on the origin of one item rather than on the history of a language highlights a methodological difference, which can be captured in the notion of the etymological approach. Describing the history of a language or even a word – in the tradition of *histoire des mots* (see Mailhammer 2007: 142–143) – assumes a perspective forward in time, from the point of origin to the present. By contrast, etymology as the investigation of the point of origin assumes a perspective backwards in time, a “retrograde” perspective (Mailhammer 2007: 143). Its purpose is not to explain the present directly through history but indirectly, by showing how and when something originated. A historical study would be focused on the question “how did this word arrive at this particular form and meaning?”, with associated implications about formal and semantic changes that make statements about the history of a language. For instance, the formal changes that have been captured in the formulation of the High German sound shift and the semantic developments since the earliest attestations would be foregrounded. By contrast, an etymological study in the spirit of the “etymological approach” is concerned first and foremost with the question “what is the origin of a word?”. That both perspectives are complementary and may in fact incorporate the same information is correct but irrelevant in this connection, because the purpose of the investigation is different. A historical investigation aims at finding out about the present or a development from a chronologically earlier to a chronologically later stage of a language, whereas an etymological investigation primarily is about getting to the point of where something comes into being. Of course, the historical perspective should incor-

porate this moment as starting point of the history, but etymology in this narrow sense is not concerned with the more recent history, but only with a point in time, it is in fact possibly even ahistorical. Etymological research starts from the earliest attestation and works its way back in time (if necessary), whereas a historical investigation works its way forwards from the point that etymological research established.

Thus the etymological approach can be defined as a quest for the origin of something rather than its history. It can be extended beyond linguistics. For instance, a linguistic etymology of the English verb *bear* will give the information that this word was formed in Proto-Indo-European as a thematic present with the meaning ‘carry’, but it is less concerned with subsequent changes. In a similar vein, a non-linguistic etymological investigation may ask about the origin of a historical event, such as the Second World War, or a political entity, such as the British Empire, without concern for subsequent developments. Of course, the etymological approach feeds on historical information, but its goal is different from a historical approach. This is why etymology is both the backbone of historical linguistics and an application at the same time.

3 Lexical and structural etymology

Traditionally, etymology has been mainly perceived as being concerned with words, but there is no reason why this should necessarily be the case. In principle anything can be studied etymologically, simply by asking the question, “where did that come from?”. The items investigated may be a phoneme or an entire text, and thus the question might be considered part of different disciplines, not even linguistic ones in the case of texts. Thus, one might wonder where affricates in High German dialects come from, what the origin of the 3sg present tense marker of Modern English full verbs is, what the origin of Fr *guerre* is, or when the plot of Shakespeare's *The Merchant of Venice* originated. As a result, it seems unproblematic to extend the notion of linguistic etymology and etymological research beyond words and word histories to anything that can be linguistically described. From a historical perspective, linguistic units with a relatively high time stability may appear more interesting than more short-lived entities that can be created spontaneously, such as sentences. But this is clearly a matter of personal taste rather than of principle.

As suggested in its title, this volume explicitly attempts to go beyond the more traditional notion of etymology as focusing on words and their history in extending the etymological approach to linguistic units outside the classical lexical domain. In two senses these extensions can be seen as structural rather

than lexical. First, constructions, such as relative clauses, are defined by their structure, and the origin of such a structure is the object of a subfield of etymology that can be called “structural etymology”. Second, structural components of language, for instance, grammatical morphemes, such as pronouns, TAM markers, and periphrastic forms, e.g. the English passive, can be investigated etymologically by applying the same research strategy. Many of the studies assembled in this volume follow one of these two pathways beyond classical lexical etymology and apply the etymological approach to structures or structural units.

To sum up, the etymological approach can be applied to any linguistic entity, phonemes, lexical items, morphemes and phrasal expressions and other phraseological units and even sentences (e.g. as constructions in the sense of e.g. Construction Grammar, see Croft 2001 and Hendery, this volume). It is the focus on the origin that defines the etymological perspective rather than the subject of investigation. However, for the sake of systematicity and clarity a subdivision of etymology into lexical and structural etymology is advocated in this book.²

4 The studies in this book

The etymology of this book – in the sense of the concept developed in the preceding sections – is a symposium at the Kioloa campus of the Australian National University in April 2010, the *First Kioloa Symposium on Etymology*. The idea was to bring together researchers from vastly different linguistic backgrounds, language focus and linguistic orientation but united in their interest in investigating the origin of phenomena they encountered in the languages they studied. This mix spawned interesting discussions about pathways of investigation, about methodologies, and about the origins and the genesis of words and structures in very different languages ranging from Indo-European to Austronesian and Australian languages.

All contributions to this volume underwent a peer review process. A substantial amount were presented at the *First Kioloa Symposium on Etymology*, though some were solicited by invitation. In two cases presenters submitted papers different to their conference papers, and one contribution had to be withdrawn just before the volume went to press. As a result, this book comprises eleven chapters, each of which illustrates the etymological approach from different viewpoints and with different phenomena in different languages.

² To my knowledge, the terms “structural etymology” and “lexical etymology” were coined by Theo Vennemann (2000); see discussion in Mailhammer (2007: ch. 3).

The chapter by the editor opens the volume outlining a framework for proposing loan etymologies that aims at removing a frequently observed bias against such etymologies by proposing a transparent and rigorous set of guidelines. This takes up a little-noticed message from a seminal paper by Milroy and Milroy (1985), which is that what is “normal transmission” of a language depends greatly on the situation, and cannot always be equated with trans-generational transmission within a homogenous community. There are situations in which a contact-influenced transmission is actually more probable, and this needs to be reflected in etymological investigation.

In a similar vein, chapter two, by Harold Koch and Luise Hercus, proposes a break with tradition in an effort to achieve a breakthrough in etymological research in Australian languages, which requires a greatly increased pool of cognates. Reworking a principle by Oswald Szemerényi, Koch and Hercus develop the point that the reflexes of an established protoform may not necessarily be attested with the same semantics or formal characteristics. Chances are that there are cognates that appear in disguised shape or with a shifted meaning, and incorporating these into etymological equations results in a much bigger number of cognates. Accordingly, the task is not to find only the semantically equivalent and transparent cognates but all cognates.

The third paper (by Harold Koch) is an application of the etymological approach focusing on the etymology of the 3sg feminine pronoun in the Pama-Nyungan languages, the largest Australian language family. Koch illustrates not only his method of morphological reconstruction (Koch 1996 *et passim*), but also showcases etymology going beyond words in uncovering the intricate shifts in meaning of grammatical forms and the formal developments of morphological markers from syntactic constructions and the absorption of morphological elements into lexical stems. One important methodological question Koch’s paper raises is how many cognates and in what distribution across a family are needed in order to “nail” an etymology, i.e. to demonstrate that the forms under comparison are more likely to reflect common ancestry than borrowing or chance.

In her contribution, Rachel Hendery presents her take on structural etymology. She uses a construction-grammar based approach to trace and typologize developmental paths of relative clauses in a range of languages. In the spirit of sections 2 and 3 above, she focuses explicitly on the origin of constructions in an etymological sense.

Chapter five is situated at the juncture between lexical and structural etymology, as it tracks down the origin of a word as well as that of a paradigm beyond what is common in etymology, namely across different languages. Theo Vennemann homes in on the history of the English reflexive pronoun, *myself*, *yourself*, etc., for which recent literature has suggested an origin under Celtic influence.

Now, for Vennemann this is not the end of the story, because one has to wonder why Celtic expressed the reflexive in such a way. It turns out that among the Indo-European languages only English and Celtic share this kind of reflexive, and Vennemann proposes that in Celtic too, it is the result of language contact, namely with Phoenician-Punic. This is an illustration of what Vennemann (2002) called the “transitivity of language contact”. It shows that etymological research does not stop within one language, but that in the case of a contact etymology one has to pursue the matter further.

Chapter six, by Patrick McConvell, illustrates the key point made by Koch and Hercus, namely that cognates can survive with drastically different semantics, in this case the shift from ‘mother’s mother’ to ‘father’s mother’. In Australian kinship systems the conceptual difference between the patriline and the matriline is highly significant, so such a shift is by no means trivial. That it can happen nevertheless demonstrates the validity of Koch and Hercus’s etymological research principle.

The next two papers, by Alexandre François and Meredith Osmond respectively, investigate pathways of semantic change in order to reveal something about words connected to the spirit world in Oceanic languages. François assembles an array of languages in detail in order to systematize avenues of semantic change in this semantic domain. Similarly, but within the limit of a much smaller paper, Osmond picks up a suggestion by Robert Blust on the origin of spirit marker terms in Oceanic languages and ties it together from an Austronesian perspective.

In his contribution, John Giacon, assembles a detailed summary of etymologies of bird names in a group of Australian languages, and gives a rich appendix illustrating the need of philologically fine-grained sifting through the evidence in order to arrive at sound conclusions about the origin of linguistic items.

The next chapter by David Nash zooms in on the name of one particular bird, the budgerigar. In minute detail Nash shows the path to discovering how this word came into English, positing that it is a case of loan blend rather than the borrowing of a word from one language.

The final paper in the volume is by Luise Hercus. Delving into her intimate knowledge of Arabana-Wangkangurru country, she shows how cognates have survived in place names in non-transparent form, but as valuable evidence in historical research, thus showing another successful application of the etymological research principle developed in chapter two.

The studies assembled here demonstrate the essence of the etymological approach and its relationship to historical linguistics, as outlined in sections 1 to 3 above. They show that a question about the origin of something is bigger than the etymology of a word, that in fact it is an investigative perspective in

its own right, and that it is in a symbiotic relationship with the question about why languages are the way they are, which crucially goes beyond a mere description, and can frequently only be answered from a historical viewpoint.

5 Acknowledgments

This book would not have been possible without the kind and efficient help of a great many people. I am very grateful to the series editors, Cynthia Allen, Harold Koch and Malcolm Ross, for opening up the series *Studies in Language Change*, and to everyone at De Gruyter Mouton, especially Emily Farrell, Wolfgang Konwitschny and Uri Tadmor, for their patience and all their help. My heartfelt gratitude also goes out to the reviewers, who did a great job in helping to ensure the academic quality of the contributions. Last, but not least, I would like to thank all the contributors for the insights they offer to the linguistic community, and for their patience and co-operation. Finally, thanks also to my family for putting up with the many hours behind the desk and for their moral support.

References

- Croft, William. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: OUP.
- Koch, Harold. 1996. Reconstruction in Morphology. In: Mark Durie & Malcolm Ross (eds.), *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*, 218–263. New York & Oxford: Oxford University Press.
- Mailhammer, Robert. 2007. *The Germanic Strong Verbs: Foundations and Development of a New System* (Trends in Linguistics Studies and Monographs 183). Berlin/New York: Mouton de Gruyter.
- Milroy, James & Leslie Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal for Linguistics* 21. 339–384.
- Seebold, Elmar. 1980. Etymologie und Lautgesetz. In: Manfred Mayrhofer, Manfred Peters & Oskar E. Pfeiffer (eds.), *Lautgeschichte und Etymologie: Akten der VI. Fachtagung der Indogermanischen Gesellschaft, Wien, 24.–29. September 1978*, 431–484. Wiesbaden: Dr. Ludwig Reichert.
- Vennemann, Theo. 2000. Zur Entstehung des Germanischen. *Sprachwissenschaft* 25. 233–269.
- Vennemann, Theo. 2002. Semitic → Celtic → English: The transitivity of language contact. In: Markku Filppula, Juhani Klemola & Heli Pitkänen (eds.), *The Celtic Roots of English* (Studies in Languages 37), 295–330. Joensuu: University of Joensuu, Faculty of Humanities.

- McConvell, Patrick & Mary Laughren. 2004. The Ngumpin-Yapa subgroup. In: Claire Bowern & Harold Koch (eds.), *Australian languages: Classification and the comparative method*, 151–177. Amsterdam & Philadelphia: John Benjamins.
- McDonald, Maryalyce & Stephen A. Wurm. 1979. *Basic materials in Wanġumara (Galali): Grammar, sentences and vocabulary* (Pacific Linguistics B-65). Canberra: Australian National University.
- Morphy, Frances. 1983. Djapu, a Yolngu dialect. In: R. M. W. Dixon & B. J. Blake (eds), *Handbook of Australian languages*. Vol. 3. xxiv–188. Canberra: Australian National University Press.
- Nash, David. 1979. *Preliminary vocabulary of the Warlmanpa language*. Cambridge, MA: Massachusetts Institute of Technology unpublished manuscript.
- Nash, David. 1996. Pronominal clitic variation in the Yapa languages: some historical speculations. In: William McGregor (ed.), *Studies in Kimberley languages in honour of Howard Coate*, 117–138. Munich: Lincom Europa.
- Nash, David. 2011. Dawes' Law generalised: Cluster simplification in the coastal dialect of the Sydney Language. In: Brett Baker, Ilana Mushin, Mark Harvey & Rod Gardner (eds), *Indigenous language and social identity: Papers in honour of Michael Walsh* (Pacific Linguistics 626), 159–178. Canberra: Australian National University.
- O'Grady, Geoff & Hale, Kenneth. 2004. The coherence and distinctiveness of the Pama-Nyungan language family within the Australian linguistic phylum. In: Claire Bowern & Harold Koch (eds.), *Australian languages: Classification and the comparative method*, 69–92. Amsterdam & Philadelphia: John Benjamins.
- Schmidt, Wilhelm. 1919. *Die Gliederung der australischen Sprachen*. Vienna: Meharisten-Buchdruckerei.
- Sharpe, Margaret. 1998. *Dictionary of Yugambah including neighbouring dialects* (Pacific Linguistics C-139). Canberra: Australian National University.
- Sharpe, Margaret C. with chapters by Marjorie Oakes, Terry Crowley, Jill Fraser-Knowles. 1996. *An introduction to the Yugambah-Bundjalung language and its dialects*. 3rd edn. Armidale, N.S.W.: University of New England.
- Smythe, W. E. 1978. Bandjalang grammar. In: Terry Crowley (ed.), *The Middle Clarence dialects of Bandjalang*, 247–478. Canberra: Australian Institute of Aboriginal Studies.
- Sutton, Peter & Harold Koch. 2008. Australian languages: A singular vision (Review article on R. M. W. Dixon 2002, Australian languages: their nature and development). *Journal of Linguistics* 44. 471–504.
- Terrill, Angela. 1998. *Biri*. (Languages of the World/Materials 258) Munich: Lincom Europa.
- Tsunoda, Tasaku. 1981. *The Djaru language of Kimberley, Western Australia* (Pacific Linguistics B-78). Canberra: Australian National University.
- Walsh, M. J. & Stephen A. Wurm. 1981. Maps of Australia and Tasmania. In: Stephen A. Wurm & S. Hattori (eds.), *Language atlas of the Pacific area. Part 1: New Guinea area, Oceania, Australia*, Maps 20–23. Canberra: Australian Academy of the Humanities.
- Waters, Bruce E. 1989. *Djinang and Djinpa: A grammatical and historical perspective* (Pacific Linguistics C-114). Canberra: Australian National University.
- Wurm, Stephen A. 1972. *Languages of Australia and Tasmania*. The Hague: Mouton.

Rachel Hendery

Constructional etymology: The sources of relative clauses¹

1 Introduction

The etymological relationship between relative clauses and other constructions is complex, because a diachronic connection between grammatical markers in two different constructions does not straightforwardly map onto an equivalent etymological relationship between the two constructions as a whole. For one thing, the route by which a marker extends into another construction can be circuitous. For another, differences between the two constructions may remain, and the two constructions may persist side-by-side in the language, stably filling overlapping functions. This is quite different from the source–outcome model that remains at the core of most grammatical research.

In this paper I discuss the diachronic relationship between relative clauses and similar constructions and argue that rather than an “etymological” source–outcome relationship, the diachronic interaction between these is often more of a matter of analogical change, frequently with participation of more than two constructions. The mechanisms involved are similar to those that lead to transfer of a relative clause marker from one language to another, in cases of constructional borrowing or calquing. In the process, I show that the sources of relative clause constructions are more varied than has previously been claimed in the typological literature. As well as the well-known sources of relative pronouns, such as demonstratives and interrogatives, other lexical and grammatical elements such as classifiers, generic nouns, discourse markers and personal pronouns can take on relative clause marking functions.

Relative clauses provide an accessible case study in constructional etymology, firstly because their synchronic typology is well understood, and secondly because relative constructions in individual languages are usually included in grammatical descriptions, so that we have well attested examples of relative clauses from multiple historical stages of many languages, including non-Indo-

¹ The data for this paper were collected as part of the basis for Hendery (2012), an extensive description of the diachronic typology of relative clauses. The examples included in that study are from published grammars and papers, and, as cannot be avoided in broad typological studies, I rely on the author's own analysis of the constructions in question when I do not have prior knowledge of the language. The full details of the constructions and languages take into account can be found in the appendices to Hendery (2012).

European languages. I would argue that the etymologies of relative clause constructions, as discussed in this paper, have implications for our understanding of constructional etymologies in general. Specifically, these findings suggest that we need to avoid overgeneralising and reducing constructional change to a simple path with a single starting point and single endpoint. Instead, constructions need to be viewed as participating in a web of other constructions, idioms, and lexemes that share similar functions, features, or forms. These other constructions that need to be taken into account may not even be all from a single language, as multilingualism and other forms of language contact are extremely widespread phenomena.

One of the reasons why discussion of constructional etymology often does not take such a wide view of potential and actual “inputs” or sources is that it is rarely a subject of theoretical or typological discussion. Rather, constructional etymology tends to be shoehorned into the concept of “grammaticalization”, or mentioned only as a secondary consideration to the synchronic analysis of the constructions in question. Changes in constructions or from one construction to another are sometimes drawn on as support for a particular syntactic model (e.g. Kikuta 2002, who presents a syntactic analysis of the Japanese internally headed relative clause, argued for on the basis of the facts of its historical development). Under the grammaticalization model, discussion has mainly focused on individual grammatical items, e.g. relative clause markers, or verb forms: only rarely entire constructions. This is natural, as grammaticalization is usually seen as a change affecting lexical items, so the pattern of an entire clause is difficult to subsume under this phenomenon (or group of co-occurring phenomena.)

A little attention has also been paid to constructional etymologies within the framework of construction grammar. Croft discusses the diachrony of constructions briefly in section 3.3 of *Radical Construction Grammar* (2001: 126). He refers to constructions diverging syntactically through replacement, renewal or split, and acquiring new semantic functions. The specific path involved, according to Croft, is that a construction extends to a new function, competes with the construction previously fulfilling this function, the latter loses, and the new one changes formally in some way that distinguishes it from itself in its former function (2001: 126). This is a very narrow model of constructional change that does not take into account the role of multiple influences on the diachronic process. Nor does it allow for the possibility that the “end point” of the process may be a stable ongoing competition between the old construction and the new, rather than the disappearance of the former (cf. Hendery 2012: 195), or, in fact, that the new construction might lose the competition and retreat back to its single former function.

Noël (2007) engages with the concept of constructional etymology within a discussion of why and how construction grammar and grammaticalization theorists should engage with each other. He also argues that the initial formation of a construction is a very different process from its subsequent change, and subject to different processes. The data examined in Hendery (2012) and this paper do not support this argument. Noël argues that schematization is the way a construction initially forms. Schematization is where a formerly loose association of features becomes used together frequently enough in a certain context that the very structure becomes associated with a specific meaning, which then transfers to the individual lexical, morphological or structural elements themselves. While this process no doubt does take place, the more frequent way that a construction enters a language where it formerly did not exist seems to be by calquing, analogical formation, or simple adaptation and extension of another construction. Perhaps these latter should not be considered true cases of new construction formation, but in that case the creation of a truly new construction is extremely rare, and not a process that we usually need to consider when looking for the etymology of any given construction.

Fried (2009) takes a more nuanced view of the diachrony of constructions. She suggests that unlike a lexical element, the etymology of a construction needs to be considered on three levels: the internal level (i.e. each individual feature), the level of the whole construction, and the discourse level. I would suggest that it is the interplay between these levels, and the fact that each level can be influenced by other constructions or languages independently of the others, that makes constructional etymology so multilayered and difficult to capture.

Work on constructional etymology has otherwise generally been from the point of view of examining individual constructions in single languages, rather than taking a typological perspective or attempting a synthetic theoretical analysis. Such studies of individual languages formed the basis of the data that were synthesized for Hendery (2012), and for this paper.

In what follows I will summarize the variety of inputs that are known to contribute to relative clause constructions cross-linguistically, and will show how complex these inputs and their interactions can be. I will relate these case studies to the wider question of how to approach the task of discovering, understanding and representing constructional etymologies. Section 2 outlines the variety of ways in which two constructions can participate in etymological relationships with each other. Section 3 discusses the question of distinguishing etymological relationships between markers of constructions from the relationships between those constructions as a whole. Having established these conceptual matters, Section 4 then turns to the etymologies of relative clauses specifically.

Finally, in Section 5, I suggest some ways of visualising these complex relationships, and the advantages of representing constructional etymologies like this.

2 “Sources” of constructions

If etymology is “the application, at the level of an individual word, of methods and insights drawn from many different areas of historical linguistics, in order to produce a coherent account of that word’s history” (Durkin 2009: 2), it seems natural to extend this concept and to talk about the *etymology of constructions*. On closer examination, we find that sometimes the “methods” and “insights”, the “areas” of linguistics, and the type of “coherent account” that we need for a construction can be quite different from those that we need for an individual lexeme. It is these sorts of differences that I will discuss in this section.

While lexical etymologies sometimes involve analogy, reanalysis, grammatical or other complex mechanisms of change, the more “prototypical” sort of etymology relates one form-meaning pair to another, with the difference between the two usually consisting of a combination of sound change and some (usually fairly transparent) semantic shift. For example Modern English *friend* goes back to Old English *frēond* ‘friend’, historically a nominalized present participle belonging to Proto-Germanic **frijōjanan* ‘to love’ (reflected in OE *frēogan* ‘to love, to favour’).² Another English example is *window* from Old Norse *vindauga*, ‘wind eye,’ replacing older words which are examples of similar compounding etymologies: *ēagþyrl*, literally ‘eye-hole,’ and *ēagduru*, literally ‘eye-door’, and eventually replacing the French-derived *fenester*. An example involving more obvious borrowing is *sky*, borrowed from Old Norse *sky* ‘cloud’, eventually replacing *heofon* > *heaven* in the sense of the ‘upper regions of the air’. This sort of etymology presumes that there was a continuous chain of speakers from a continuously-existing speech community using form-meaning pairs that gradually changed from A_1 to A_n via A_2 A_3 A_4 , etc.³

² All etymologies in this paragraph are taken from the *Oxford English Dictionary*, online edition. November 2010. Oxford University Press. Sighted 16 January 2011.

³ While I refer to these relatively simple etymologies as ‘prototypical’, I do not mean to imply that lexical etymologies are always this straightforward. Robert Mailhammer (p.c.) points out that even lexical etymologies can involve multiple sources, as in the case of blends, calques or taboo replacements. However, it seems to me that in the few cases where constructional etymology is discussed in the literature in a theoretical sense, the assumption is that rather straightforward ‘source-outcome’ relationships exist: a model reflected more in the most basic sorts of (lexical) etymology.

This prototypical sort of etymology can also be involved in constructional change.⁴ A construction C_1 is a form-function pair where the form consists of a word order pattern, fixed morphological or lexical marking, and a syntactic relationship to other phrases or clauses in the sentence. Changes in any of these and/or extension of the construction to a new function can create a new construction C_2 , which we could say is etymologically related to C_1 . We would want to describe C_1 and C_n this way if the changes in form and function were gradual such that a continuous chain of etymologically related constructions $C_1, C_2, C_3, \dots, C_n$ had existed. This sort of change is schematized in Figure 1.

Straightforward source—outcome relationship

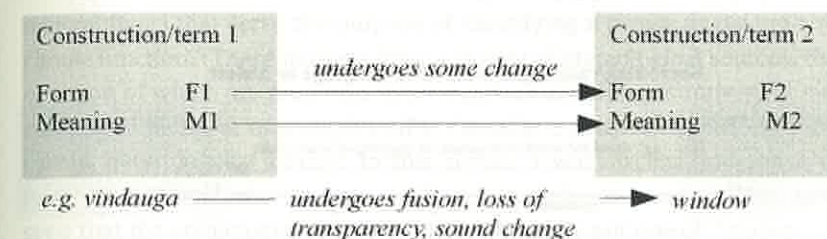


Figure 1: Schematization of a source-outcome relationship

However, as I will show in the later sections of this paper, with relative clauses at least, this source–outcome model seems almost never to be the case. Instead, we find different sources for the changes in form and the changes in function, or abrupt borrowing or calquing, so that C_1 and C_2 are not related in this sort of gradual etymology. These sorts of changes are schematized in figures 2 and 3.

In the case of analogical change, the “source” continues to exist alongside the “outcome”. In the case of borrowing or calquing, the source continues to exist in the donor language or dialect, or if it does not, its disappearance is not directly connected to the fact it was the source of the new construction. Moreover, F2 is usually wildly different from F1 because it is frequently a calque, or in any case heavily adapted to the recipient language. The morphological forms, lexical forms, or word order may be different (although not usually all three of these – otherwise we would not identify the construction as based on that of the

⁴ See also Mailhammer (2007: chapter 3) for proposing etymology as a research perspective that can be applied to lexical as well as to structural items (Lexical vs. Structural Etymology, see Vennemann 2000), which can easily be extended to constructions.

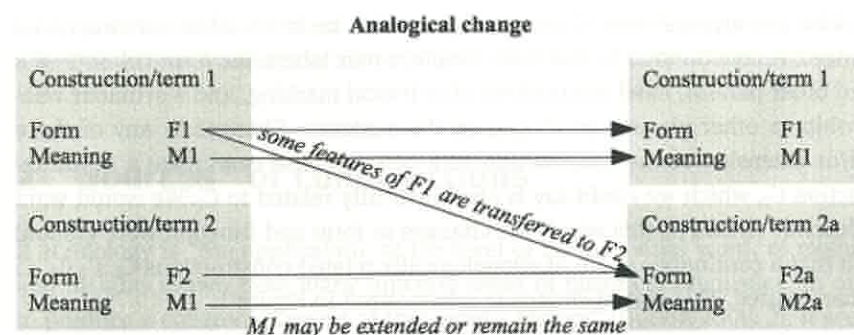


Figure 2: Schematization of analogical change

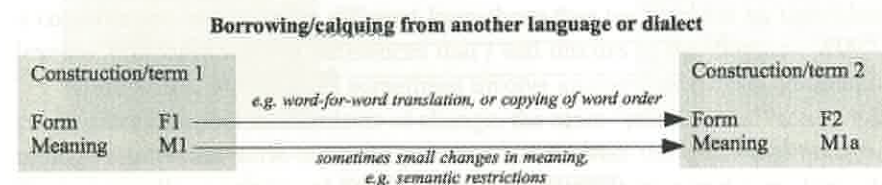


Figure 3: Schematization of borrowing or calquing

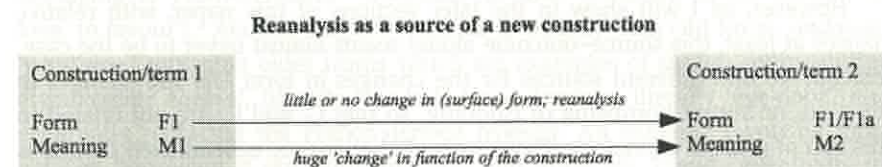


Figure 4: Schematization of reanalysis

donor language at all). M1a on the other hand is usually identical or extremely similar to M1 (which is why I have labelled it M1a instead of M2).

Finally, reanalysis can lead to a new construction or term, but again this is different from the sort of scenario represented in figure 1, since construction 1 often continues to exist once construction 2 arises.

With this outline of the various possibilities in mind, I will now move on to a discussion of the actual sources of a particular construction, the relative clause.

3 Relationship between sources of constructions and sources of lexemes that mark that construction

In order to understand the etymologies or sources of relative clause constructions, we can begin with the simpler question of the sources of relative clause markers. While I will argue that the relationship between these two sorts of etymologies is not a straightforward equation, one nonetheless provides a clue to the other.

Relative clause markers are defined for the purpose of this paper as grammatical items that serve the purpose of identifying a clause as having a relative clause function.⁵ These markers have a variety of etymological sources, the most common of which are pronouns (interrogatives and demonstratives) or elements commonly found in non-relative NPs: classifiers, possessive markers, “generic” nouns, even definite articles. In this section I will discuss how these various types of markers came to be used in relative clause constructions. It will be seen that the etymological relationship is generally not one of “grammaticalization” but nor is it as straight-forward as the sort of development discussed at the start of section two.

The simplest and by far most common approach to the etymology of relative clause markers is to look for obviously cognate forms in other constructions: demonstratives, interrogatives, and other clausal constructions such as NP and VP complementation, adverbial clauses, etc. This is the approach taken, for example, by Heine and Kuteva (2002). Under the assumptions of “grammaticalization”, such markers are thought to have gradually become part of the relative clause construction through such mechanisms as phonological reduction, fusion, and semantic bleaching. In fact, however, these are strange mechanisms to assume for the development of relative clause markers, which, unlike the prototypical cases of grammaticalization, are most frequently already grammatical items when they are co-opted for use in relative clauses. A survey of known sources of relative clause-marking elements is still worthwhile, but as an end in itself it is rather uninteresting, because it does not tell us how the etymological change proceeds. For this we have to turn to case studies where a change has happened during a period of a language for which we have written attestation or otherwise strong evidence for the directionality of the change and the steps it took on the way.

⁵ In other words, I am subsuming relative pronouns, relative complementizers, and relative verb affixes under one category here.

In Table 1 I list the sources of relative clause markers that are best documented as historical developments. These are not simply based on apparent cognates, but are cases where the directionality of the change is well-evidenced, usually because the development occurred during a period of the language with written history. More details on these developments can be found in Hendery (2012). Other possible sources that are less well documented include comparatives, possessive pronouns, and indefinite pronouns.

Source of RC marker	Evidenced in
interrogatives	many languages
demonstratives	many languages
discourse markers (e.g. topic or focus markers)	Quechua, Basque, Tok Pisin, PIE, and possibly many others
generic nouns	Malay, Lhasa Tibetan, Newari, Japanese, Hebrew, Korean
classifiers	Thai, various Amazonian languages, Newari, Cantonese
personal pronouns	Gothic

Table 1: Sources of relative clause markers

From a comparison of these case studies, it can be seen that development of new relative clause markers generally proceeds in one of two different ways, the first of which is through reanalysis: fossilization of an element that happens to appear in relative clauses or on the clause boundary, but is not originally part of the relative clause construction per se. This is the case for some generic nouns, classifiers, and various discourse markers, and can be illustrated with the case of Tibetan, as outlined in DeLancey (1986). Among various other relative clause markers, Lhasa Tibetan has one, *sa*, that literally means ‘earth’ or ‘place’, while Newari has *mha*, meaning ‘body’. DeLancey proposes that these may have been ‘kidnapped’ into the relative clause from an appositive construction, something like example (1).

- (1) [[[*jĩ-ĩ* *nyan-a*]-*mha*] [*nya*]
 [NP [CP I-ERG buy-PART]-live.thing] [NP fish]
 ‘the fish, the live thing I bought.’

If the clause boundaries on example 1 were reanalysed, the resulting construction is the modern-day Newari relative clause illustrated in example (2).

- (2) [[[*jĩ-ĩ* *nyan-a* -*mha*] *nya*]
 [NP [CP I-ERG buy-PART-live.thing] fish]
 ‘the fish, the live thing I bought.’

A second way for relative clause markers to arise is through analogy with other constructions that resemble the relative clause, but even the clearest examples of this pathway, examined in Hendery (2012), are ambiguous in terms of directionality. Possessive constructions, for example, resemble relative clauses in many languages, and in Japanese (*no*), Malay (*punya*) and the Semitic languages (various markers), there is enough historical evidence to trace the development. The problems with these examples are that modern Japanese *no* does not function as a relative clause marker, except in child language; Malay *punya* is not yet unambiguously a relative clause marker, although it seems to be moving in that direction, and the Semitic particles are all related to demonstratives as well, so that both the relative and possessive function might have independently developed from this third source. Certainly as a general rule, markers can be extended to new constructions by analogy, and Hendery (2012) shows how this is frequently the case with the extension of relative clause markers to new constructions (adverbial clauses, general subordination, adjective phrases, among others). That the reverse is not strongly evidenced is probably merely an accident of the historical record.

A few sources, such as interrogatives, appear to have primarily spread through language contact, which can also be viewed as a sort of analogical process, as the first speakers to bring them into the new language are presumably doing so on analogy with the relative clause constructions in their other language. Interrogatives in colloquial Basque relative clauses, for example, are thought to have developed on analogy with the French and Spanish constructions (Hurch 1989). A language contact cause has also been suggested for English interrogative relative clauses (with French as the source), and is also thought to have been the case in Nahuatl (Hill and Hill 1981), Quechua (Lefebvre and Muysken 1982: 45; Karttunen 2000: 403), and the Uralic (Comrie 1988: 474–475) and Turkic languages (Johanson 1998: 333, 335) that have interrogative-based relative clause markers. It has even been suggested that this is nearly always the case when the interrogative-based relative pronoun construction is found in non-Indo-European languages (Comrie 1998: 61).

It is therefore the case that even the etymologies of relative clause markers can be quite complex. They can have more than one source, as in the Basque example, where the Basque interrogative *zein*, the French relative pronoun *qui*, etc. and the Spanish relative pronouns *que*, etc. are all “sources” (inputs) in some sense for the new Basque relative pronoun *zein*.

The sorts of elements that become relative clause markers do so partly because they are well suited to the role, in the sense that they already perform one or more of the characteristic operations involved in relativization: subordination, attribution or pronominalization (cf. Lehmann 1984: chapter 4). For example in

English, even if the demonstrative *that* were not a relative clause marker, it would still perform all of these functions: it can stand in a modifying relationship to a noun in the NP (*that man*), it can introduce a subordinate clause (*I said that he left*) and it can stand alone as a referential pronoun (*I saw that*). This sets it up perfectly to function as a relative clause marker, since it already plays all the necessary roles. In this sense, the structural features of a source construction shape the path of change taken by the marker, which is the sort of change that most closely resembles the prototypical single-source etymological path.

It is possible that some pathways of development toward relative clause markers are not unidirectional. Possessive markers, comparatives, focus markers, and indefinite pronouns are all elements which might be both source and outcome of relative clause marker change. In Hendery (2012), however, I show that none of these is well documented for both directions. For possessive markers becoming relative clause markers, there is strong evidence of direction, but for the reverse direction, the evidence is weaker and there are plausible alternative explanations. The same is true of indefinite pronouns. Focus markers are well-attested as sources of relative clauses, but not so strongly as outcomes. For comparatives, neither direction is well-documented or without alternative possible explanations. If there are markers for which both directions of change occur, however, this is unsurprising since the developments discussed here are not “grammaticalization” in the strict sense of the term, but rather are the extension of a grammatical marker from one grammatical function to another.

In this section we have seen that the etymologies of relative clause markers can resemble the more prototypical lexical etymologies, in that they can have a single source, consist of semantic and formal changes, and involve loss of the element in its original function so that it really does appear that a change has occurred. On the other hand, when relative clause markers arise through analogy, whether it is analogy with another construction in the same language, or with relative clauses in other languages of the region, their etymology is more complicated, with multiple “sources”. As we turn to the etymology of whole constructions in the next section, we will find this latter type of relationship between source and outcome constructions, and even more complex variations where different components of the construction have different source relationships.

4 Relative clauses: constructional change and etymology

As the examination of relative clause markers above suggested might be the case, the diachronic interaction between relative clauses and other constructions

is often a matter of analogical change rather than a source–outcome relationship. In cases of constructional borrowing or calquing, the mechanisms involved for whole constructions are also similar to those that lead to transfer of a relative clause marker from one language to another.

4.1 “Sources” of relative clauses

Just as is the case for relative clause markers, a cursory glance at relative clause constructions in historical perspective may lead to the impression that a simple source–outcome relationship exists. In the Ethiopian Semitic languages, for example, the modern prenominal RC seems to have developed directly out of earlier postnominal constructions. In these Ethiopian languages a drastic restructuring of the word order and phrase order rules took place, presumably due to influence of the surrounding Cushitic languages. The relative clause construction was literally flipped around the head noun as part of this reordering. It can therefore be said to have developed directly out of the older construction, which did not persist. However, one could also argue that the various Cushitic relative clause constructions were sources of the new Semitic one as well, but note that this is not a one-to-one relationship, as it is not one specific construction in a single Cushitic language that is said to be the source.

In other cases, a source–outcome relationship between two different constructions might exist, but it is only one of many relationships that the construction participates in. For example, in 12th century Middle English, a relative clause construction developed that was introduced with an interrogative pronoun. This could be said to be a new relative clause that had its etymological source in embedded questions. Certainly it resembles these closely. However, the construction also has diachronic relationships⁶ with other constructions: pre-existing relative clause constructions in English, and relative clause constructions in French. These relationships are evidenced by the fact that it resembled the pre-existing relative clauses in terms of position, word order, and other features, and even participated in blended constructions, such as *whom that* relative clauses, illustrated in example (3).

⁶ And it presumably also had synchronic relationships with these other constructions. Speakers would have seen the set of constructions involving interrogatives as ‘similar’ to each other in some way, and saw the set of constructions that served the same function (relativization) as ‘similar’ to each other in some other ways. It is this connection that led to the transfer of features among the various constructions in the first place.

- (3) *Only the sighte of hire, whom that I serve, ... wolde han suffised right ynough for me*

'Only the sight of her, whom (that) I serve, ... would have sufficed right enough for me' (Knight's Tale 1231: line 373–5)

The adoption of interrogative-marked relative clauses may also have been triggered or accelerated by the fact that this "constructional polysemy" already existed in the Norman French spoken by many residents of England at that time, as illustrated in example (4). It is a development which may have occurred anyway, however, as *wh*-pronouns were already occurring in the heads of "free" relative clauses, as in example (5), which meant a simple reanalysis could have turned them into relative pronouns even without an exact model.

- (4) *Irieement parla li lous, ki mult esteit cuntrarious*
Angrily spoke the wolf, who very was quarrelsome
(*Lais of Marie de France* 2: 5)

- (5) *Sodes is þe sylle swa hwæt swa [þu me byddest]*
Truly I you give so **what-acc** as [you me ask]
'Truly I give you whatever you ask of me' (*Gospel of St Mark in the West-Saxon Gospels* 290, from Allen 1980: 278, ex. 51).

Part of the problem with thinking of constructions in terms of the source–outcome etymological model is that there are usually multiple types and subtypes of a construction in a single language. Therefore, the construction already participates in a kind of network of other constructions, which are treated by speakers as similar to varying degrees. This means that changes can enter a construction from multiple sources in this network. (For example in the Middle English example, a new construction was created with the function of one, the markers of another, word order and other restrictions from several more, and on analogy with yet another construction that existed in the grammars of some bilingual speakers.)

Note that the connections in Figure 5 are not meant to represent causal relationships, or one-way source–outcome connections. Rather they represent relationships of *any sort*, including source–outcome relationships, but also analogies, constructions that were both simultaneously present in a speaker's grammar and resemble each other, constructions that participated in blends, or competed with each other (e.g. one was extended into the other's domain). This is the field of constructions that may have participated in the etymology of the Middle English interrogative-marked relative clause. The diagram shows how

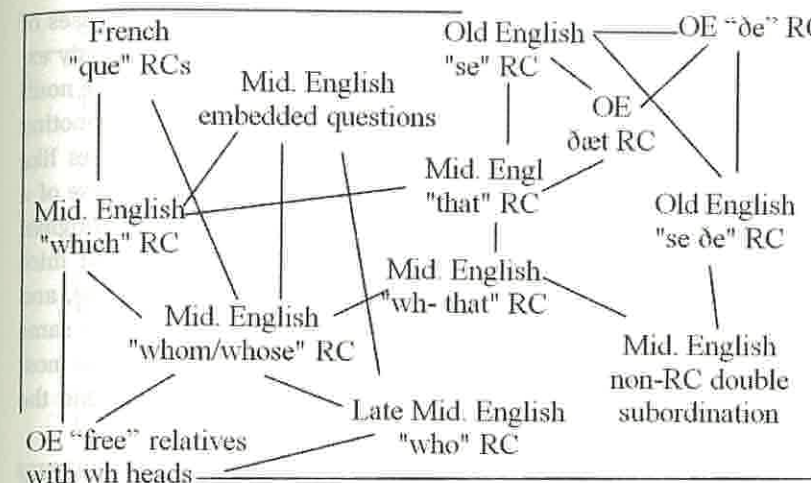


Figure 5: Middle English interrogative-based relative clauses

much of the complexity of the context is lost when we try to reduce it to a simple one-to-one source–outcome relationship.

Because multiple types of construction with the same or similar function exist, sometimes nothing is lost when a new construction is created. The new construction can co-exist side-by-side with other pre-existing constructions that fulfil the same function, and with its source or sources. This was the case with the English interrogative-based relative clauses: the construction with *that* persisted alongside the new alternatives. Although the distribution of each type is not identical, there are some contexts in which either can occur. This means that it does not really make sense to talk about it as a "change" in any one relative clause construction: the interrogative relative clause is not a change to the demonstrative-marked construction. Rather these sorts of changes are changes to a *system*, through a sort of (non-identical) duplication of the form of a construction, sometimes with extension to a new function.

This is unlike the simple kind of etymology illustrated by *friend*, as detailed in section two, since in that case the related verb was lost, and the morphology as a present participle became opaque. The example of ME *wind-oge* 'wind-eye' > *window*, is also different, because in that case the 'wind-eye' term with its transparent meaning and exact morphology and phonological form is (gradually) lost and replaced by the new single-morpheme form. It is also unlike the sort of etymology illustrated by ONorse *sky* 'cloud' > MnE *sky*, where the old meaning is lost and replaced by the new meaning to which it was extended.

Rather it is more similar to the sorts of etymologies exhibited in some cases of zero derivation: English *to mouse* 'to use a (computer) mouse' was recently extended to this new meaning/function, while the language retains both the noun *mouse*, meaning a type of computer peripheral, and the original word denoting the animal. Grammatically and morphologically the new *mouse* behaves like other verbs (*mousing*, *moused*). Interestingly, the noun *mouse* in the sense of a computer peripheral seems to be distinct in speakers' minds from the original *mouse*, as, for at least some speakers, the two have different plurals: *mice* versus (*computer*) *mouses*. The difference between this particular example, and our examples of relative clause "etymologies" is that a verb with the same meaning as *to mouse* did not previously exist in the language, while most languages already have RC constructions before a new one develops, and the old construction often continues to exist alongside the new one.⁷

In Hendery (2012) I categorized a large number of relative clause changes into types, which are summarized in Table 2.⁸ None of these types proved to involve simple, direct changes of one construction into another. In no case can we say that a given construction has *an* etymological source. In order to illustrate this etymological complexity, in the rest of this section I will discuss examples of each of these types of change.

4.1.1 Extension (a)

When a relative clause marker is extended to new types of construction, this essentially creates a new relative clause. The new construction has the syntactic form of one older construction, and the morphological marking of another. This makes it difficult to say exactly which older construction is the etymological source of the new one. Such cases are common, but the history of Russian provides a particularly clear illustration.

Russian has several different relative clause markers which have changed their distribution. The element *izhe*, originally inflecting for gender, number

⁷ Sometimes when two or more relative clauses continue to co-exist in a language, they each become used in different contexts, i.e. one will become most commonly used for animate nouns, another for inanimates, and/or one will be used for definite nouns, and one for indefinites, and/or one for subject relative clauses and another for objects. These sorts of developments of complementary or near-complementary distribution have occurred in the Romance languages, English, and Russian, among others. See Hendery (2012: 133–134 onwards) for further details.

⁸ The languages listed in the right-hand column are not an exhaustive list, but they are those for which there is the most evidence of change and information about what exactly occurred.

Change	Languages affected
a. Extension (and/or restriction) of the 'types' of RC a marker can be used in	Many languages, including English, Quechua, Turkish, the Celtic languages, Russian, the Romance languages, etc.
b. Grammatical change in category of the RC marker (e.g. affixation, loss or development of inflection)	English, Celtic, Bantu languages, Akkadian, Russian, Egyptian.
c. Development of a RC with a 'balanced' verb construction, where previously only deranked constructions existed	Japanese, Hebrew, Turkish, Hungarian, Finnish, various Mongolian languages, possibly Egyptian, possibly proto-Indo-European.
d. Development of a RC with a deranked verb construction, where previously only balanced constructions existed	Armenian, possibly Egyptian, possibly proto-Indo-European (we do not know the directionality of the change in the latter two)
e. Development of a prenominal RC where only postnominals, correlatives or internally headed RCs had previously existed	Armenian, perhaps various other branches of Indo-European (pre-historic)
f. Development of a correlative RC where one had previously not existed in the language	Some L2-based New Englishes, possibly the Dravidian languages.
g. Development of a postnominal RC where these had previously not existed in the language	Quechua, Basque, perhaps Turkish, Telugu, and most branches of Indo-European (pre-historic).

Table 2: Structural changes in relative clauses

and case, but later used invariantly, was restricted to postnominal, embedded relative clauses, but was later extended to correlatives as well (Leckey 1992: 16–17, 21). The element *kotoryi*, which inflects for number, gender and case, underwent the opposite expansion. It was originally used only in correlatives, but then extended to postnominal embedded relative clauses (Leckey 1992: 253, 259–260).⁹

Leckey suggests that this was an analogical development, meaning that in some sense the two different relative clause constructions were *both* sources of each of the two "new" constructions:

The increasing frequency of the preposed RC with *izhe* (from the time period attested by Ickler to that attested here) was an analogical development. As *kotoryi* began functioning in environments typical for *izhe*, *izhe* became eligible for environments typical of *kotoryi*. (Leckey 1992: 259–260)

⁹ The two elements were not in competition for very long, as *izhe* was then lost altogether.

The diagram in figure 6 schematizes this development.

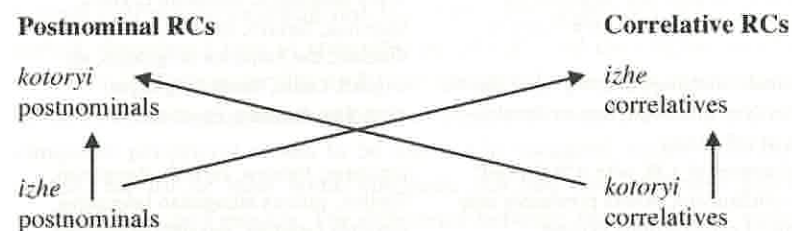


Figure 6: Medieval Russian relative clauses. (based on Leckey 1992)

4.1.2 Change in category of relative marker (b)

The change of grammatical category of a relative clause marker also creates a new construction. Rather than a clause introduced by a relative pronoun, the new construction might be marked with an invariant subordinator, or even with a special verb form, in cases where the relative clause marker cliticized and then fused to the verb (as is thought to have occurred in Nguni, for example (Zeller 2004).)

In English the relative marker *that*, while not originally a relative pronoun, is becoming more pronominal in current English usage (Seppänen and Kjellmer 1995; Seppänen 1997). A genitive form *that's* has come into usage in some varieties. In English it is clearly an analogy between *who* – *whose* leading to *that* – *that's*. It is hard to imagine that such a development would occur in a language that did not already have a pre-existing inflecting relative pronoun. We can say that the construction in (6a) has at least three “sources”, which are given in (6b–d). The first source is the construction that is actually being replaced in the dialect of the speaker of (6a). The second is the standard *that* construction, which is contributing the marker to the new construction. And the third is the *wh*-relative-clause, where there is minimally a distinction between possessive and non-possessive forms, and which is therefore an analogical source for the new construction.

- (6) a. *The table that's leg is broken*
 b. *The table, the leg of which is broken*
 c. *The table that I painted*
 d. *The boy whose leg is broken*

4.1.3 Development of a new “balanced” relative clause (c)

The development of a new “balanced” relative clause construction can be illustrated by Turkish.

Turkish has a standard deranked relative clause construction, but also a minor secondary strategy involving the marker *ki* and a balanced postnominal relative clause. In some dialects this has become the main relative clause construction (Johanson 1998: 333, 335). There are at least two sources of this new construction. The first is the fact that some verbs in the deranked paradigm happen to be identical to the corresponding main clause forms (Antinucci, Duranti, and Gebert 1979: 162–163). This means that some standard deranked relative clauses looked like balanced constructions, and the new, *truly* balanced construction may have had its roots in an extension of these. The second source, however, is thought to have been heavy contact with Iranian and/or Slavic languages, as the dialects in which the balanced construction became standard are those in closest contact with these models. The relative clause marker *ki* itself is from Persian *ke* (Johanson 1998: 335).

4.1.4 Development of a deranked prenominal relative clause (d and e)

Armenian has taken the opposite route, and developed a deranked, prenominal relative clause, where it previously had only had postnominal and correlative constructions. In Classical Armenian, relative clauses were expressed with postnominal or correlative constructions (Hewitt 1978: 100–113). In Modern Armenian it is also possible to use prenominal relative constructions with deranked verb forms: participles, otherwise used as gerunds (Hewitt 1978: 128–129, fn. 20). These three types of relative clause are illustrated in example (7).

- (7) a. *əntreaç erkotasans, z-or-s ew aʔakeals anouaneaç*
 he-chose twelve whom indeed apostles he-named
 ‘he chose twelve, whom he named apostles’ (Classical Armenian, Luke 6: 13, in Hewitt 1978: 102).
- b. *or oğ I noçanē val aʔis haşşē, zna*
 who ever from them quickly to me reaches-SUBJ, him
tağawor arariç
 king I-shall-make-SUBJ.FUT
 ‘Whichever of them reaches me the sooner, him shall I make king’
 (Classical Armenian, Eznik, in Hewitt 1978: 105).

- c. *çourtēn tandž-ou-ol ajs xelč mardə*
 by-the-cold-ABL being-tormented this poor man-the
 'this poor man who is being tormented by the cold...'
 (Modern Western Armenian, Feydit 1948: 127 in Hewitt 1978: 124)

It is usually thought that the deranked prenominal construction is a calque on the Turkish relative clause (Feydit 1948: 352) or due to influence from literary Russian (Stilo 2004: 46), but Hewitt argues that it could have developed in a series of gradual changes from the pre-existing postnominal construction. The scenario he outlines is that the auxiliary verb in a sentence like "the church which was built by Gregory the Great..." was lost, and then the relative clause marker was also eliminated, since the clause began to be reinterpreted as phrasal. This creates a construction like the English participial relative clause: *the church built by Gregory the Great...* These intermediate constructions are attested in Classical Armenian. Finally, the participle, now analogically seen as an adjective, moved to the same position in which adjectives in Armenian are most common: prenominally (Hewitt 1978: 128).

Since we know that Armenian has been heavily influenced by Turkish, it is probable that both Hewitt's suggested path of change *and* language contact were involved: that Armenian recreated the Turkish relative clause strategy using resources already available to it. The question of whether it was a language-internal development hastened and encouraged by outside influence, or a language contact phenomenon supported by internal typological developments is simply a matter of perspective. The deranked prenominal Armenian relative clause therefore may have had three sources: the Armenian postnominal relative clause, the adjective phrase (an analogical source), and the Turkish relative clause.

4.1.5 Development of a correlative construction (f)

The development of a correlative relative clause is less commonly attested than other developments. There are no documented cases of the gradual, organic evolution of a correlative construction out of any other type of relative clause. What is known to occur, however, is the transfer of correlative patterns from L1 to L2 when speakers acquire a new language. In at least one case, South African Indian English, this has led to the widespread use of correlatives in a newly created dialect (Mesthrie and Dunne 1990). An example of this is given in example (8).

- (8) *Which-one I put in the jar, that-one is good*
 (Mesthrie and Dunne 1990: 37, ex. 13)

The use of an interrogative pronoun suggests some relationship to the Standard English relative clause. Hindi, the "donor" language for this construction¹⁰, does not use an interrogative here, but rather a specifically relative pronoun. So once again, there are at least two sources contributing to this new construction.

4.1.6 Development of a new postnominal construction (g)

Finally, to illustrate the development of a new postnominal relative clause construction, I will take the example of Quechua. In Quechua postnominal relative clauses such as the one illustrated in example 9 are assumed to be modelled on Spanish. These are clearly a recent development, as according to Lefebvre and Muysken (1982: 45) they are not described in the otherwise detailed early grammars such as that by Holguín ([1609] 1842), or that by Ricardo ([1586] 1970). Nor are they found in the conservative modern dialects of the language, but only in Cuzco Quechua.

- (9) *warma-wan puklla-ra-ni [pi-wan-mi Joseca rima-sqa-n]*
 girl-with play-PAST-1 [who-with-af Jose talk-PAST-3]
 'I played with the girl that Jose talked with' (Lefebvre and Muysken 1982: 3, ex. 3)

This example contrasts with the relative clause constructions that are more widely used among the various Quechua dialects, i.e. internally headed relative clauses, prenominal and correlative RCs with nominalized verbs (Lefebvre and Muysken 1982; Cole 1985; Grosu 1998). A correlative construction with an interrogative-based relative marker that looks a little like some sort of intermediate stage between the older and more recent relative clause types also exists and is illustrated in example (10).

- (10) *pi-ta-n khuya-nki chay-ta-n ñoqa-pas khuya-ni*
 who-acc-af love-2 that-acc-af I-too love-1
 'who you love, I love too'. (Ricardo, 1970)

¹⁰ South African Indian English speakers may have either a northern or southern Indian background. Hindi is the original L1 of the speakers with a northern background, and it seems to be members of this subset of the community that use the correlative construction. Those with a southern background, speakers of Dravidian languages, use a Dravidian-style participial relative clause instead (Mesthrie and Dunne 1990: 35–38).

This construction was attested as early as 1586, before the postnominal construction appears. This suggests that the interrogative-based relative pronoun (possibly calqued from Spanish relative clauses) developed first in the correlative construction, and then was extended or reborrowed into a new construction that was a more direct copy of the Spanish one. Even those who argue that the interrogative-marked construction is a Quechua-internal development, not a calque (e.g. Lefebvre 1984; Appel and Muysken 2006: 161), admit that the postnominal position of the new construction is at the very least most likely a result of word order changes triggered by Spanish influence (cf. Appel and Muysken 2006: 161). So once again, it seems that the etymology of the new construction is complicated: both the traditional Quechua correlative construction and the Spanish relative clause have fed into the new postnominal RC.

In this section we have seen that for each of the types of new relative clause a language can develop, the “etymology” is complex, with multiple sources playing a role. The examples chosen here were not deliberately selected because they illustrate this point, but only because they are the best-documented cases of the various changes, so we can be most certain of the historical details. In any of the examples in Table 2 for which we have sufficient information to enable close examination, we find the same complexity.

4.2 Non-RC constructions as sources of relative clauses

In the previous section, most of the sources of new relative clause constructions discussed were other types of relative clause construction in the language, and relative clauses in other languages with which speakers were in contact. In this section I will discuss the few non-relative constructions that are found as sources of relative clauses. The constructions in question include simple noun phrases, possessives, adverbial clauses, and comparatives. Once again, it will be seen that the new construction is rarely the result of one simple change operating on a single source.

4.2.1 Simple noun phrases

Simple noun phrases frequently serve as points of analogical extension to the relative clause. Classifiers, demonstratives, and possibly articles are all elements whose prototypical function belongs to the simple noun phrase, but can be adopted as relative clause markers. Even then, however, the simple noun phrase itself is not the sole source of the new relative clause construction. Take, for

example, the origin of relative clause constructions that are marked with classifier morphemes. Such constructions are found in various Amazonian languages (cf. Aikhenvald 2000: 88–94); Newari (DeLancey 1986), Cantonese (Shi and Li 2002); and Thai (Sornhiran 1978). An example of an Amazonian classifier-marked relative clause construction (from the language Tuyuca) is shown in example (11).

- (11) *ti- bā- ré ādō- pé*
 that- CL:PATH- RE here- THEM.CONTR
 [*kā atí- a- ri- bā- pi*]
 [3MSG come- recently- SG.NOM- CL:PATH- LOC]
hoá- wa- yigí
 start.down.path- go- 3MSG.PAST.EVIDENTIAL
 ‘He started down that path over here [that he had recently come on]’
 (Aikhenvald 2000: 93)

Aikhenvald (2000: 88) suggests that the characteristic of classifiers that leads them to develop into relative clause markers may be their ability to function anaphorically. An example of this anaphoric use can be seen in the example from the Akateko language in example (12).

- (12) *yeešin ši nax tšonwom š?ey tšotan nax*
 all.right said CL:MAN merchant sat down CL:MAN
smaxa šyetšmane nax sat k'al k'am tšen tumin
 waiting waited CL:MAN long.time there.was.no CL:ROCK money
 ‘All right, said the merchant and sat down to wait, he waited, he waited a long time, but there was no money.’ (Aikhenvald 2000: 88)

It seems, in fact, that all the examples of languages in which classifiers appear to have become relative clause markers also demonstrate a similar anaphoric use of classifiers. So these anaphoric classifier constructions may well be another source contributing to the evolution of the classifier-relative-clause construction.

It is also the case that classifiers occur with a number of other components as well as in a simple noun phrase. While the most well known classifiers occur with numerals, in some languages there are also classifiers that accompany possessives and/or verbs. Cantonese is a language that illustrates the wide range of contexts in which such classifier elements can appear. Historically, the Cantonese element *ge* had both a demonstrative use and a classifier use (origi-

nally denoting lengths of bamboo, later extended to arrows, then candles and certain animals, then to fruit, birds and people, and now used as a general “default” classifier (Aikhenvald 2000: 401). It also came to be used in possessives, adjective phrases, adverbial phrases and to link demonstratives to the corresponding noun (Matthews and Yip 1994: 88–90). But because the order of its historical spread into the various functions is unclear (Shi and Li 2002), we cannot be certain whether the relative clause construction with *ge* was based on one or more of these other constructions, on its demonstrative use, or original classifier use. At the very least, both the demonstrative and numeral classifier functions of *ge* already existed when the relative clause construction developed, so it is possible that these both played a role in the evolution of the latter. If more of the other *ge* constructions already existed, in particular the possessive, adjective and adverbial “linking” functions, it is possible that *ge* simply became reanalysed as a general-use linking particle, and the new relative clause construction was created on this basis.

Classifier constructions are therefore good examples of cases where multiple other constructions also seem to serve as sources for a new relative clause.

4.2.2 Possessives

Although possessive constructions and relative clauses share markers and other features in many languages, it is hard to pin down cases where one is undeniably the source of the other. Generally there are other equally plausible explanations for the similarity, or a third type of construction provides an intermediate step between the two. Japanese is a good illustration of this messiness. In Japanese since the 16th century, a headless relative clause construction has existed, which is exemplified in example (13). This construction is discussed in depth in Yap, Matthews, and Horie (2004), who call it an “S-pronominal construction”.

- (13) *Taroo-ga kat-ta no takakat-ta*
 TAROO-NOM buy-PAST NO expensive-PAST
 ‘the one Taro bought was expensive’ (Yap, Matthews, and Horie 2004: 142, ex. 12)

In the late 19th century until the early 20th century, it was possible to use a similar construction for headed relative clauses. This sort of construction still occurs in child language, so it is clearly a natural extension of example (13). But where did (13) come from?

There are several theories about this. Yap, Matthews, and Horie (2004) suggest that the *no* in this construction is ultimately the same *no* used as a possessive

marker in Japanese since the 8th century. Martin (1991) has an alternative theory: that this *no* was originally *mono* ‘thing’, which explains why it was first used only in headless relative clauses. No matter which of these is the ultimate source, both agree that the relative clause construction was mediated by at least one other construction: what Yap, Matthews, and Horie call the “N-pronominal”. This construction is shown in example (14), and can be used in such ellipses as we would use the corresponding ‘s in English: e.g. *John’s is over there or Whose is it? Sarah’s*. Martin (1991: 282–283) suggests that this *no* here, too, is a reduction of *mono* ‘thing’.

- (4) *Taroo no*
 Taroo no
 ‘Taroo’s’ (N-pronominal function (Yap, Matthews, and Horie 2004: 141 ex. 7))

In Yap, Matthews, and Horie’s explanation, then, we have both the possessive construction (*Taroo no hon* ‘Taroo’s book’) and the ‘N-pronominal’ function both feeding into the creation of a new (headless) relative clause construction. In Martin’s scenario, only the construction with (*mo*)*no* in the N-pronominal construction is the source of the relative clause. Since *no* existed at the same time in a possessive use, however, as soon as *mono* was reduced to *no*, speakers would have been able to see the possessive and relative constructions as related. The evidence that they did so, or at least that they lost the connection with the *mono* construction is the extension of the *no* relative clause to headed constructions in the late 19th century.

4.2.3 Adverbial clauses

The relationship between adverbial clauses and relative clauses is a little more complicated. This is because extension seems to occur in either direction (cf. Hendery 2012: chapter 2). It is also because a boundary between true adverbial clauses and true relative clauses is not straightforward to draw: adverbial clauses introduced with *where* in English, for example, are simply headless relative clauses, and the same holds for the locative adverbial marker in many other languages. In this section I will summarize the development of the *pou*-marked relative clause construction in Greek as presented by Nicholas (1998), since this is perhaps the best-documented case available.

From the earliest attested period where relative clauses with (*ho*)*pou* appear in Greek, (*ho*)*pou* is also found in adverbial clauses. Nicholas (1998: 178) argues

that this is because (*ho*)*pou* was created on analogy with the pre-existing *hostis* and therefore it inherited the multiple functions of this element, including indefinite headless relativization and embedded questions. The indefinite headless relatives (which, when locative, are adverbial clauses), and the “true” relatives are compared in examples (15)–(18).

Headless indefinite RC with *hostis*

- (15) *dó:so: gâr díph^hron te dúo:*
Give.FUT.IS for chariot.ACC and two
t^heriaók^henas híppous...
and arching-neck.MASC.ACC.PL HORSE.MASC.ACC.PL
hóstis ke tlaíe: hoí t^hautó:I
REL.M.NOM.S particle dare.OPT.PRES.3S 3S.DAT and'self.3MS
kú:dos ároito.
glory.ACC win.AOR.OPT.MID.3S

‘For I will give him a chariot, and two horses of arching neck, ...

to **whosoever** shall dare the deed, and for himself shall win glory.’

(Iliad X 305–7, via Nicholas 1998: 172 ex. 8b)

Adverbial clause (locative headless indefinite RC) with *hopou*

- (16) *hópou d^hapóllo:n skaiòs ê:I*
Where for'Apollo ignorant.NOM.MS be.3S.PRES
tínes sop^hoí?
anyone.M.PL.NOM wise.M.PL.NOM?
‘Where Apollo is ignorant shall men be wise?’
(Euripides Electra 972 via Nicholas 1998: 178 ex. 11a)

True relative clause with *hostis*

- (17) *áp^hron dè: keínós ge kai*
foolish.MS.NOM indeed DEM.MS.NOM at.least and
outidanòs pélei ané:r, hóstis
worthless.MS.NOM be.3S.PRES man.MS.NOM REL
kseinodóko:I érida prop^hére:tai aét^hlo:n
host.MS.DAT strife.FS.ACC propose.3S.PRES.SUBJ contest.MPL.GEN

‘That man indeed is foolish and worth nothing, whoever proposes a strife in contests with his host’

(Od VIII 210, via Nicholas 1998: 173, ex. 8c)

True relative clause with *hopou*

- (18) *eíta en tē:I toiaúte:I pólei hópou mē:*
then in a.FS.DAT such.FS.DAT state.FS.DAT where not
lógo:I érgo:I te hikanoí
thought.MS.DAT work.NS.DAT and sufficing.MPL.NOM
p^hulakes eien...
guardian.MPL.NOM be.3PL.PRES.OPT

‘In such a state, where there are no custodians competent in act as in thought [...]’

(Plato, Laws. 964C, via Nicholas 1998: 178, ex. 11c)

According to Monteil (1963: 387, cited in Nicholas 1998), at this stage *hostis* was also just starting to appear in some examples that are interpretable as specific (restrictive) relative clauses, such as the example in (17). The new form *hopou*, formed from *ho-* (article/demonstrative) and *-pou* (locative interrogative), just as *hostis* was formed from *ho-* and *-tis* (nominative interrogative), took on the same functions with a locative (and hence adverbial) flavour: marking of indirect questions about location, generalising locative adverb (‘wherever’), and marker of oblique relative clauses (Nicholas 1998: 178).

Nicholas, and before him, Tzartanos (1991) reconstructed a subsequent pathway of development that led to *pou* being used in more abstract spatial relative clauses/adverbials, adverbials of cause and then temporal adverbials. The extension of *pou* to core relatives (subject and direct object NPreIs) occurred in Late Middle Greek or Early Modern Greek. There are three clear examples from the fifth, sixth and seventh centuries (Nicholas 1998: 207), and then none at all until 980 AD. The fifth century example is given in (19).

Non-locative RC with *hopou* (Middle Greek)

- (19) *tòn adelphòn hópou eíkhe pròs tē:n*
the.MS.ACC brother.MS.ACC REL have.3S towards the.FS.ACC.
autòn lúpe:n
him.MS.ACC sorrow.FS.ACC
‘the brother who had distressed him’
(Aphrodisiaca Patrum 300B, via Nicholas 1998: 207, ex. 34a)

The Greek data provide a demonstration of how constructional polysemy can arise without gradual extension. The *pou* construction was polysemous from the start, because *pou* was created on analogy with another element which *had*

undergone extension and inherited the polysemy of its source. The etymological sources of the *pou*-marked relative clauses in Greek are therefore simultaneously but to varying degrees: indirect questions, adverbial clauses, generalising indefinite (headless) relative clauses, articles/demonstratives, and interrogatives. These relationships are depicted in the diagram in Figure 7.

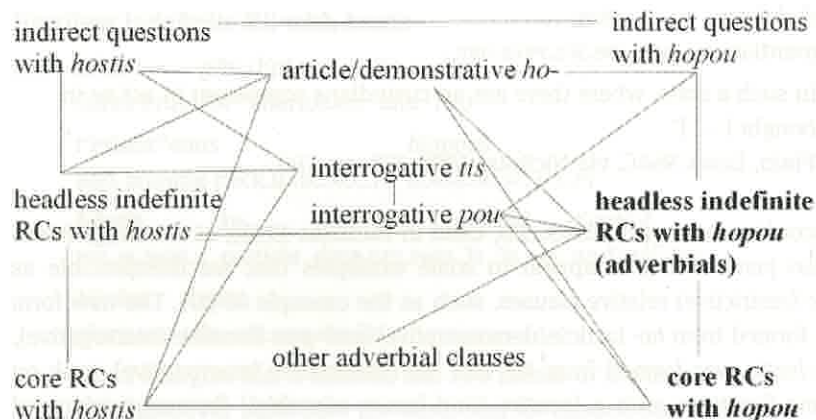


Figure 7: Sources contributing to Greek relative clauses with (ho)pou

4.2.4 Comparatives

The diachronic relationship between comparative and relative constructions most frequently seems to be a matter of relative clauses being a source of comparatives (Haspelmath and Buchholz 1998: 290), but the other direction probably occurred in English and Old Norse (Faarlund 2004: 259). I will take Norse as an example here.

The first occurrence of the comparative marker *sem* in a Norse relative clause construction was in an inscription where it was combined with another marker *es*, shown here in example 20. This construction therefore looks like a blend between the comparative and a pre-existing relative clause: so once again we have at least two sources contributing to the construction's etymology.

- (20) *sa hit aki sims (=sem es) uti furs*
 This was.called Aake, sims (=REL REL) abroad traveled.REFL
 'This was Aake, who traveled abroad' (=died) (Lindblad 1943: 113)

As well as a primary comparative function, the marker *sem* was also found in adverbials of place, which may have provided a bridging construction to the relative clause in some contexts.

A pre-existing relative clause construction in Norse, marked with *er*, was also multifunctional, and could be used in comparative constructions such as example 21. This means it was almost certainly a model for the new *sem* construction.

- (21) *pá fundu menn hans í gamma einum*
 then found.3PL men.NOM his in hut.DAT one.M.DAT
konu pá er þeir höfðu enga eét jamvæna
 woman.ACC that REL they.M had.3PL not.F.ACC seen even-beautiful.F.ACC
 'Then his men found in a hut a woman so beautiful that they had never seen anyone like her' (Faarlund 2004: 261)

Finally, it is thought that *sem* developed relative clause marking functions first in the Norwegian dialects of Old Norse, spreading later to Iceland, Sweden and then Denmark (Noreen 1923: 319; Faarlund 2004: 259). In these latter varieties, the comparative construction with *sem* existed when the relative construction was adopted, so we could say there were three sources contributing to the new construction in these cases: the *sem*-comparative/adverbial, the pre-existing relative clause (with *er*), and the 'foreign' *sem* relative clause in the dialect from which it spread.

In the previous section we saw that relative clause constructions that develop out of *other* relative constructions had complex, multi-faceted etymologies, and in this section we have seen that the same holds for relative clauses that develop out of non-relative constructions. In fact, even describing the changes as one construction "developing out of" one another is misleading. It is better to visualize a construction as participating in a web of diachronic interactions – and the way in which we carve up this web and declare everything in one space to be one construction and everything in another to be a different one, is fairly arbitrary. This becomes very clear when we try to represent constructional etymologies visually. In the following section I will discuss the ways in which constructional etymologies are sometimes represented visually and will suggest that these representations can be improved if we do not gloss over the sorts of complex relationships that I have discussed above.

5 Conceptual space representations

I have shown in the previous sections how inclusion of multiple sources, potential sources, and language contact influences in constructional etymologies leads to a more accurate understanding of how and why relative clauses end

up looking the way they do. One difficulty with these more complex etymologies, however, is that they are much more difficult to represent visually (and two-dimensionally) than simpler source-outcome models. In this section I will begin by introducing the sorts of representations of constructional change that are frequently found in the literature, and will then move on to presenting some suggestions for how they can be extended to represent the more complex etymologies that I have discussed above.

Diagrams of semantic or functional change are common in the grammatical literature, but sometimes consist of rather linear, unidirectional pathways, where each function is a step mediating between a source and an outcome, as in the example from Van der Auwera and Plungian (1998: 88) reproduced in Figure 8.

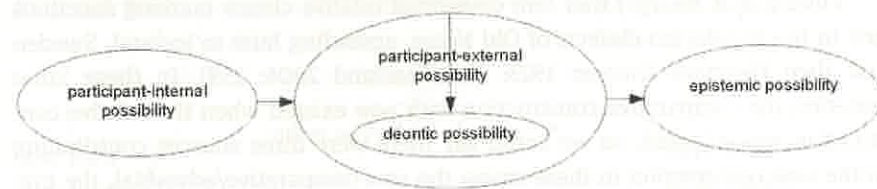


Figure 8: Grammaticalization of possibilities, from Van der Auwera and Plungian (1998: 88)

The diachronic semantic map in Figure 8 shows the “grammaticalization pathways” of words that encode possibility. (The “deontic possibility” set is located within the set of “participant-external possibility” because it is seen as a subtype of the latter.) Van der Auwera and Plungian (1998) do also use more complex maps, with many-to-one and one-to-many connections. However, these connections are not meant to suggest multiple sources, but rather alternative sources or outcomes of change in different languages. Van der Auwera and Plungian’s diachronic maps, like many others in the literature, are intended to present a cross-linguistic summary of all possible pathways of change.

Diagrams intended to represent typological reality rather than the history of an individual language sometimes have bidirectional links, representing the fact that change in either direction is possible. This is the case for the map of nominalizers/pronominalizers in Yap, Matthews, and Horie (2004: Fig. 3), reproduced here as Figure 9. The connections between each step of the pathway are bidirectional.

Yap, Matthews, and Horie’s diagram still glosses over the fact that one construction does not smoothly “become” another. In this map that is not a serious problem, as each change is of approximately the same type: simple extension of a construction to a new function. This diagram leaves out, of course, any con-

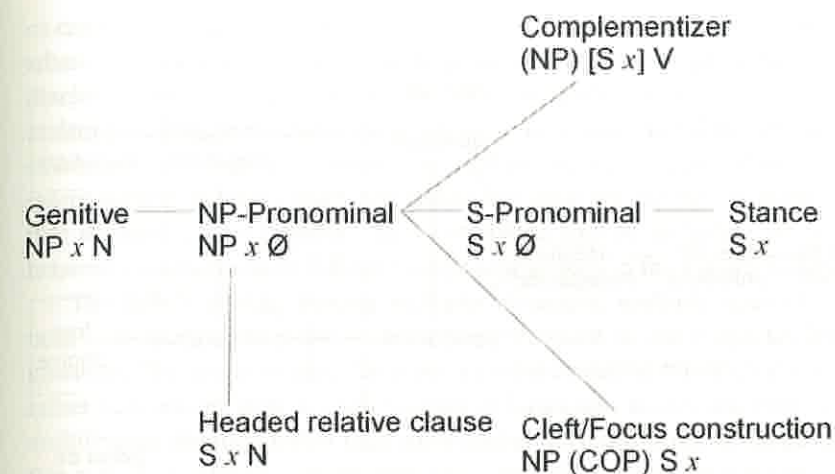


Figure 9: “Grammatical functions centered on (pro)nominalizers as an implicational map. Note that x represents the (pro)nominalizer morpheme under investigation (e.g. Japanese *no*, Mandarin *de*, and Malay *empunya*).” (From “From pronominalizer to pragmatic marker: implications for unidirectionality from a crosslinguistic perspective”. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and Down the Cline – The Nature of Grammaticalization*, 2004, pp 137–168. With kind permission by John Benjamins Publishing Company, Amsterdam/Philadelphia. www.benjamins.com)

nections to pre-existing constructions filling those functions, as well as any other models for the functional polysemy that may have existed. To represent the full story of the etymology of some relative clause constructions we would need to include not only these complicating factors, but also different kinds and different strengths of links between them and their sources. For example, we may have borrowing of a marker from one language to another, calquing without lexical borrowing, reanalysis, extension of a marker by analogy, or transfer of word order patterns.

Semantic maps and “conceptual spaces”, both for individual lexemes and for constructions, allow the mapping of the terms/constructions of an individual language onto the typological diagram, showing at the same time synchronic and diachronic reality. Croft (2001: 92–95; 2010) discusses these sorts of maps. An example for Romanian indefinite pronouns overlaid onto Haspelmath’s semantic map of indefiniteness is reproduced in Figure 10.

Croft argues that such maps should be data-driven: derived from comparison of multiple exemplars. In other words, the Romanian and other language data should first be graphed, to show which senses/functions are expressed uniquely, and which are expressed by polysemous forms. When sufficient lan-

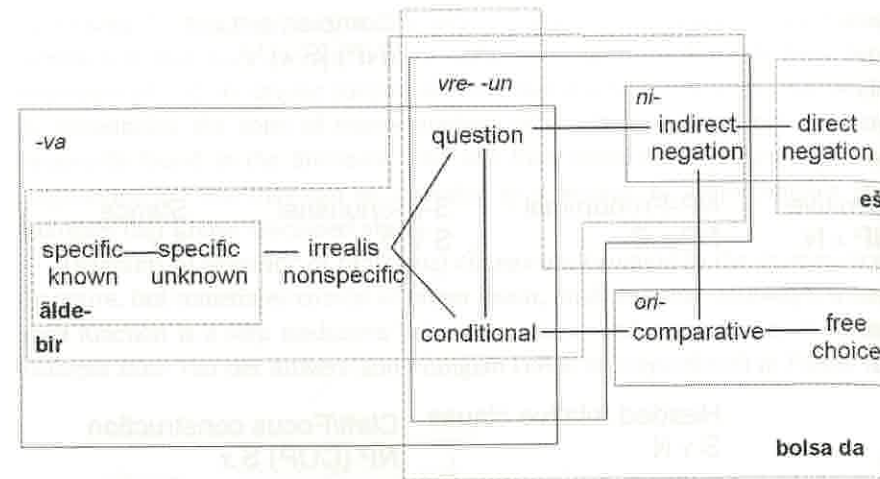


Figure 10: Croft's mapping of Romanian and Kazakh indefinite pronouns onto Haspelmath's semantic map (Croft 2010: Figure 2)

languages are included, this allows the automatic generation of a semantic map such as Haspelmath's, which underlies Figure 10.

The advantage of semantic maps is that the overlay of several languages would allow us to see the models that exist in one language for new constructions in another. For example, if we were to create a semantic map of pronouns, we would see that interrogative and relative functions in French and Spanish are both performed by the same function, and that in older Basque this was not the case, but in modern Basque it was. To see this on one map we would of course need to treat multiple stages of a language as separate languages.

The difficulty with such maps for our purposes is that when we move beyond lexical items to entire constructions, polysemy is not the only relationship we wish to represent. In the discussion in the previous sections we have seen that numerous types of diachronic and synchronic relationship can obtain between constructions: even various different types of polysemy. Two constructions can be marked with the same marker – this is a type of polysemy. But two constructions that are both postnominal, with balanced verbs and the same marker are much more similar than two that share a marker, but do not have the same position or verb type. Even when only examining the markers themselves, while analogical extension, reanalysis and even calquing can leave evidence in the shape of polysemy, direct borrowing of a marker from one language to another does not. And even when we see polysemy in a semantic map, we cannot be certain whether that polysemy is the result of extension, reanalysis

or calquing. Semantic maps, with bidirectional relationships between elements, where necessary, are a good starting point. But the ideal representation of the diachronic relationship between constructions would also be able to represent multiple languages, multiple stages of each language, and different types and strengths of relationship between constructions. Such a diagram allows us to represent something more closely resembling the speakers' reality: the constructions that speakers of the language have considered similar, including relationships between constructions in different languages spoken in the speech community.

The difficulty with showing multiple languages, multiple types of connection, and many-to-many links on the same diagram is that it quickly becomes unwieldy. The key is to treat them not as cross-linguistic typological maps, but rather to represent only a small number of languages at any one time: a single multilingual situation rather than some sort of universal truth. As an example of this sort of "diachronic construction map", the Old English interrogative-marked relative clause(s) are represented in Figure 11.

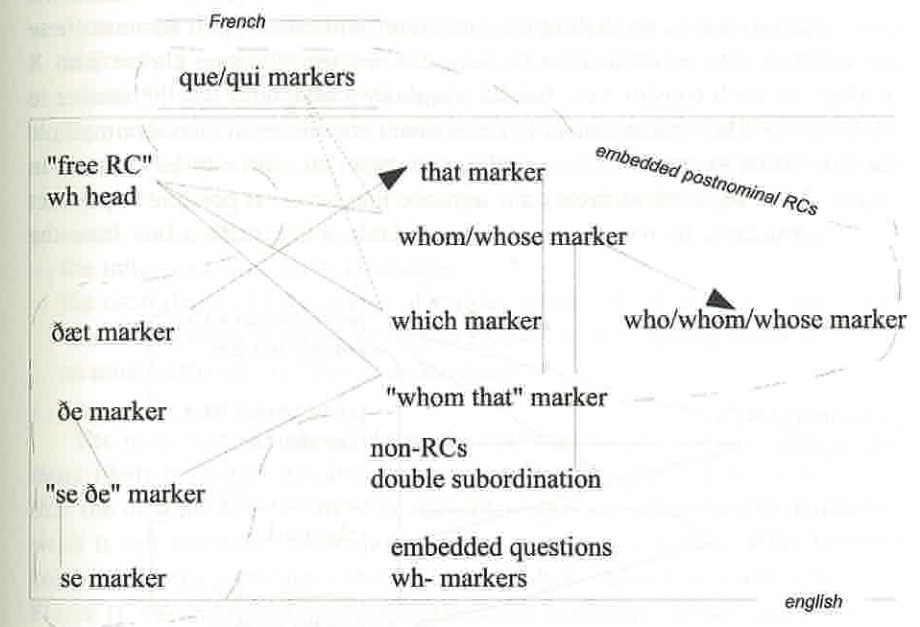


Figure 11: Old English interrogative-marked relative clauses

The two main boxes group together those constructions that belong to the same language. These boxes are labelled with the language names. The spatial arrangement of the constructions represents time: older constructions are further to the left; newer ones are further to the right. Constructions that are from the

same time period are lined up underneath each other. If desired, the columns could be labelled with the time period (Old English, Middle English and Late Middle English in this case).

Lines connecting constructions represent a formal relationship. I have used lines ending in an arrow to represent the relationship “developed directly into”. This is the case where a marker merely underwent sound change, or where one marker is just a differently case-marked version of another. Lines not ending in arrows represent a less direct relationship. This includes reanalysis, analogy and potential analogy, i.e. very similar forms occurring in similar functions, or constructions in the same time period that use the same markers for different functions. Finally, the dotted line in this case represents a potential cross-language analogy. I would suggest this could be used to represent the type of relationship that exists between e.g. two markers that are phonologically very different, but that belong to the same word class and perform the same function. This will usually be the case with calques, but may also occur within a language.

The dotted line surrounds constructions that have the same function, i.e. here: relative clauses (excluding the indefinite “whoever” type). Because these are grouped into a labelled set, I have not written “Relative clause with X marker” for each construction, but for simplicity’s sake have left the marker to represent the whole construction. In cases where one wishes to represent multiple variables such as marker, clause order, verb type, etc, this can be done as in Figure 12. Listing each variable on a separate line makes it possible to connect two constructions by means of only one variable – e.g. draw a line from the

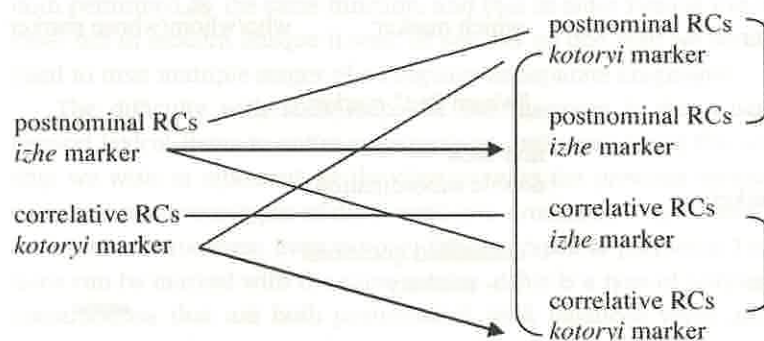


Figure 12: Russian postnominal and correlative relative clause constructions.¹¹

¹¹ In Figure 12 all the constructions of interest have the same (relative) function, and are from the same language, so the conventions I used in Figure 11 to group constructions by language or function are redundant here and have not been used.

marker in construction 1 to the marker in construction 2 to show the marker was adopted, or to draw a line from “postnominal” in construction 1 to “postnominal” in construction 2, to show that the same word order is used.

An advantage of these sorts of visual representations of constructional etymologies is that they show the difference between constructions that “develop into” another construction in a later period (in the same way that a lexical item like *frēond* ‘loved, favoured’ becomes *friend*), and constructions that spawn a new function or form, but also remain in their previous incarnation alongside the new construction. We can see this latter sort of development quite clearly in the map of Russian in Figure 12: instead of two constructions in the first column developing gradually into two other constructions in the second column, we find four constructions in the later stage of the language: two retentions and two new innovations.

It is also possible to “read off” such diachronic construction maps the following information:

- the number of possible influences on the development of the construction (how many lines connect to the construction in question? How many connect indirectly (i.e. connect to a construction that in turn connects to the one of interest?) How strong are these connections?)
- the number of other constructions in the language performing the same function (grouped within the dotted line). This has implications for how much competition the new construction has, and hence what is likely to happen next – loss, redistribution, etc.
- the influences from other languages
- the centrality or ‘fringe’ status of various constructions: are there many lines connecting them to others or only a couple? Do they belong to the same sets as most of the others? The same language?

The more tightly embedded into the “network” a given construction is, the more likely it is that speakers would consider it a “type” of the construction that the diagram focuses on. Note that this degree of embeddedness in the network is not absolute, since each diagram is only a snapshot of the network centered on one particular construction – in these cases, the relative clause. In Figure 11, the embedded question construction looks very “fringe”, but if we had instead created a “questions” map, it would no doubt be much more connected.

Finally, such diachronic construction maps provide an alternative to having to choose which variable to privilege in a taxonomy of constructions. For example, in Russian we could present a taxonomy of medieval relative clause types (based purely on those in Figure 12) as either the left-hand hierarchy or the right-hand hierarchy in Figure 13.

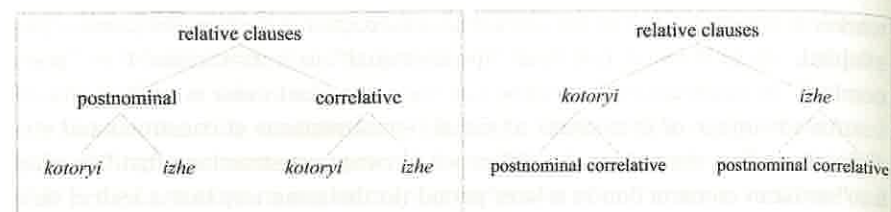


Figure 13: Two alternative taxonomies for Medieval Russian relative clauses

In the map in Figure 12, however, it is apparent that there are relationships between the late medieval constructions both in terms of the clause position and in terms of the marker, and we do not have to treat one as a subtype of the other unless we wish to. The fuzzier taxonomy is less misleading with regard to representing speakers' own intuitions about the relationship between the constructions. Unless we have psychological evidence to the contrary, there is no reason to think that speakers perceive one construction as a subtype of another, but rather they would merely be aware of (some) formal and functional similarities and differences: i.e. the relationships in the right-hand column of Figure 12.

The sorts of maps presented in this section are not ideal for all purposes. There are situations for which a hierarchical taxonomy, or a more typical semantic map or grammatical pathway diagram is more appropriate. When we want to represent a fuller picture of the diachronic situation, however, especially if language contact is a relevant factor, it is worth considering maps like those in Figure 11 or 12.

6 Conclusions

The sorts of diachronic relationships that can exist between constructions are much more complex than simple source–outcome relationships of the sort found in what we might think of as “prototypical” lexical etymologies. The types and subtypes of construction that linguists may identify are not always the same constructions that speakers treat as “similar” and “different”, yet processes of change such as analogy, borrowing or calquing are the direct result of speakers' intuitions.

As far as relative clauses specifically are concerned, I have shown here that the markers, verb types, and position of the construction are all relatively independent variables. One or more of these can be borrowed from another language, or changed on analogy with other constructions in the language. These other

constructions that can participate in analogy or other diachronic interactions include other relative clauses, but also simple noun phrases, possessives, adverbial clauses, and comparatives.

By working through examples of each of the types of attested change in relative clause constructions (see Table 2), I showed that none of these is straightforward. Even the development of a “new” construction, such as a prenominal relative clause where previously only postnominals have existed, or vice versa, is not a case of a single source construction being adapted to a new function. Rather there are frequently multiple other constructions that feed into the new one, often including constructions from another language.

For this reason, we would miss important generals by limiting a typology to a subset of constructions that share one function; nor is it ideal to look for a simple chain of constructions where the original function F1 is extended to F2, then F3, etc. To understand constructional etymology and change it is necessary to view them in the context of the whole web of other constructions that surrounds them, including those which only share a marker, or only share a function, or only share one other feature such as word order.

Such networks of constructions make for very complex typologies, which cannot be represented using current diagrammatic conventions. I therefore turned in the final section of this paper to the question of whether it is possible to capture such complex diachronic relationships visually. I outlined a suggestion for a diachronic construction map that allows a researcher to display complex and nuanced findings about synchronic relationships, language contact influences and the different sorts and strengths of diachronic relationships.

References

- Aikhenvald, Alexandra. 2000. *Classifiers: a Typology of Noun Categorization*. Oxford: Oxford University Press.
- Allen, Cynthia. 1980. *Topics in Diachronic English Syntax*. New York & London: Garland Publishing.
- Antinucci, Francesco, Alessandro Duranti & Lucyna Gebert. 1979. Relative clause structure, relative clause perception, and the change from SOV to SVO. *Cognition: International Journal of Cognitive Psychology* 7. 145–175.
- Appel, René & Pieter Muysken. 2006. *Language Contact and Bilingualism*. Amsterdam: Amsterdam University Press.
- Comrie, Bernard. 1988. General features of the Uralic languages. In: Denis Sinor (ed.), *The Uralic Languages: Description, History and Foreign Influences*, 451–477. Leiden: Brill.
- Comrie, Bernard. 1998. Rethinking the typology of relative clauses. *Language Design* 1. 59–86.
- Cristofaro, Sonia. 2005. *Subordination*. Oxford: Oxford University Press.

- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Croft, William. 2010. 'Relativity, linguistic variation and language universals', *Cognitextes* (Online) 4. <http://cognitextes.revues.org/303> (accessed 25/2/2011).
- DeLancey, Scott. 1986. Relativization as nominalization in Tibetan and Newari. Paper presented at the 19th International Conference on Sino-Tibetan Languages and Linguistics. <http://www.uoregon.edu/~delancey/papers/relnom.html> (accessed 24/2/2011).
- Durkin, Phillip. 2009. *The Oxford Guide to Etymology*, Oxford: Oxford University Press.
- Faarlund, Jan Terje. 2004. *The Syntax of Old Norse*. Oxford: Oxford University Press.
- Feydt, Frédéric. 1948. *Manuel de Langue Arménienne (Arménien Occidental Moderne)*. Paris: Klincksieck.
- Fried, Mirjam. 2009. Construction Grammar as a tool for diachronic analysis. *Constructions and Frames* 1(2). 262–291.
- Grunberg, Madéleine (ed.). 1967. *The West-Saxon Gospels*. Amsterdam: Scheltem & Holkema.
- Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hendery, Rachel. 2012. *Relative Clauses in Time and Space (Typological Studies in Language)*. Amsterdam & Philadelphia: John Benjamins.
- Hewitt, Brian. 1995. *Georgian*. Amsterdam and Philadelphia: John Benjamins.
- Hill, Jane H. & Kenneth C. Hill. 1981. Variation in relative clause construction in modern Nahuatl. In: Frances E. Karttunen (ed.), *Nahuatl Studies in Memory of Fernando Horcasitas* (Texas Linguistic Forum 18), 89–104. Austin: Dept. of Linguistics, University of Texas at Austin.
- Johanson, Lars. 1998. Code copying in Irano-Turkic. *Language Sciences* 20(3). 325–337.
- Haspelmath, Martin & Oda Buchholz. 1998. Equative and similative constructions in the languages of Europe. In: Johan van der Auwera & Dónall Baoill (eds.), *Adverbial Constructions in the Languages of Europe*, 277–334. Berlin & New York: Walter de Gruyter.
- Holguín, Diego G. 1842. [1609]. *Gramática y arte nueva de la lengua general de todo el Peru, llamada lengua quichua o lengua del Inca*. New revised and corrected ed. by Francisco de Canto, Lima.
- Hurch, Bernhard. 1989. Hispanisierung im Baskischen. In: *Vielfalt der Kontakte*, Norbert Boretzky & Birgit Iglá (eds). Essen: Brockmeyer.
- Karttunen, Frances. 2000. Language contact in Latin America. In: Thomas L. Willett (ed.), *Uto-Aztecan: Structural, Temporal & Geographic Perspectives: papers in memory of Wick R. Miller by the Friends of Uto-Aztecan*, 387–410. Hermosillo, Son: Universidad de Sonora.
- Kikuta, Chiharu Uda. 2002. Clausal complement or adverbial clause?: toward an integral account of the Japanese internally-headed relative clause. In: Frank van Eynde, Lars Hellan and Dorothee Beermann (eds), *Proceedings of the 8th International HPSG Conference, Norwegian University of Science and Technology (3–5 August 2001)*, 202–220. CSLI Publications. <http://csli-publications.stanford.edu/202> (accessed 24/2/2011).
- Leckey, Celia Elizabeth. 1992. *Relative clauses in medieval Russian texts*. Berkeley: University of California dissertation.
- Lehmann, Christian. 1984. *Der Relativsatz*. Tübingen: Gunter Narr.
- Lefebvre, Claire. 1984. Grammaires en contact: définitions et perspectives de recherche. *Revue Québécoise de Linguistique* 14(1). 11–48.
- Lefebvre, Claire & Pieter Muysken. 1982. *Relative Clauses in Cuzco Quechua: Interactions between Core and Periphery*. Bloomington: Indiana University Press.

- Lindblad, Gustaf. 1943. *Relativ Satsfögning i de Nordiska Fornspråken* [Relative clauses in the ancient Nordic languages]. Lund: Walter Ekstrands.
- Mailhammer, Robert. 2007. *The Germanic Strong Verbs: Foundations and Development of a new System*. Berlin & New York: Mouton de Gruyter.
- Martin, Mary L. (ed.). 1984. *The Fables of Marie de France*. Birmingham, AL: Summa Publications.
- Martin, Samuel E. 1991. Recent research on the relationships of Japanese and Korean. In: Sydney M. Lamb and E. Douglas Mitchell (eds.), *Sprung from Some Common Source*, 269–292. Stanford: Stanford University Press.
- Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. New York: Routledge.
- Mesthrie, Rajend & Timothy T. Dunne. 1990. Syntactic variation in language shift: The relative clause in South African Indian English. *Language Variation and Change* 2. 31–56.
- Monteil, Pierre. 1963. *La Phrase Relative en Grec Ancien*. Paris: Klincksieck.
- Nicholas, Nick. 1998. *The story of pu*. Melbourne: The University of Melbourne dissertation.
- Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14 (2). 177–202.
- Ricardo, Antonio. 1970. Reprint. *Arte y vocabulario en la lengua general del Peru llamada quichua en la lengua española*. ed. Rafael Aguilar Paez, Lima. Original edition, Lima, 1586.
- Romaine, Suzanne. 1980. The relative clause marker in Scots English: diffusion, complexity and style as dimensions of syntactic change. *Language in Society* 9. 221–249.
- Seppänen, Aimo. 1997. The genitives of the relative pronouns in present-day English. In: Jenny Cheshire and Dieter Stein (eds.), *Taming the Vernacular: From Dialect to Written Standard Language*, 152–169. London: Longman.
- Seppänen, Aimo & Göran Kjellmer. 1995. The dog that's leg was run over: on the genitive of the relative pronoun. *English Studies* 76. 389–400.
- Shi, Yuzhi & Charles N. Li. 2002. The establishment of the classifier system and the grammaticalization of the morphosyntactic particle *de* in Chinese. *Language Sciences* 24. 1–15.
- Sornhiran, Pasinhee. 1978. *A transformational study of relative clauses in Thai*. Austin: University of Texas at Austin dissertation.
- Stilo, Donald. 2004. Iranian as a buffer zone between the universal typologies of Turkic and Semitic. In: Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*, 35–63. London: Routledge-Curzon.
- Van der Auwera, Jan & Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2 (1). 79–124.
- Vennemann, Theo. 2000. Zur Entstehung des Germanischen. *Sprachwissenschaft* 25. 233–269.
- Yap, Foong Ha, Stephen Matthews & Kaoru Horie. 2004. From pronominalizer to pragmatic marker: implications for unidirectionality from a crosslinguistic perspective. In: Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and Down the Cline – The Nature of Grammaticalization*, 137–168. Amsterdam: John Benjamins.
- Zeller, Jochen. 2004. Relative clause formation in the Bantu languages of South Africa. *Southern African Linguistics and Applied Language Studies* 22 (1–2). 75–93.