

## **Workshop proposal: Getting more out of corpus data: expanding the empirical evidence for semantic change**

Organizer: Tine Breban, The University of Manchester

### Background

Since the 1980s access to historical materials has greatly improved by the arrival of historical corpora for English and empirical research into language change has soared. Over the years, more, better, and bigger corpora have been compiled. New methods for processing and analysing data from corpora have been developed. The emphasis has been on applying quantitative and statistical methods, as well as on exploring the role of the surrounding context in language change (e.g. Diewald 2002, Traugott 2012). One thing that does not appear to have changed very much since the early corpora is what we are looking for to detect language change. This is particularly a problem for changes that have been defined primarily in terms of a change in function or meaning, such as grammaticalization or subjectification. Corpora do not directly give access to meaning. It is common practice to support a hypothesized change in function and meaning with evidence of differences in form or distribution, as the latter changes can be observed empirically. Researchers have to rely on concomitant changes that can be observed in corpus data. The issue that this workshop seeks to redress is the limited types of changes in form and distribution that are usually discussed as evidence.

The problem of operationalizing semantic change has been raised by several linguists studying subjectification (Aaron and Torres Cacoulios 2005, De Smet and Verstraete 2006, Torres Cacoulios and Schwenter 2006, López-Couso 2010, Traugott 2010), and intersubjectification (Traugott 2010, Brems et al. 2014). But, perhaps more unexpectedly, it also applies to grammaticalization, which is one of the most well-studied types of semantic change in the ‘corpus age’. The foundational works on grammaticalization appeared in the 1980s-1990s, before large scale diachronic corpora were widely used. The most widely applied proposal is Lehmann’s (1995 [1982]) parameters which have become used as recognition criteria for grammaticalization in data studies. However, as is well-known, this proposal has its limitations (see e.g. Breban et al. 2012): without going into detail on issues with specific changes, the general problems are that the changes identified occur at a late stage in the grammaticalization process (Hopper 1991) and are particularly suited for languages that express grammar in a synthetic way. Empirically identifying cases of grammaticalization at an early stage for an analytic language such as English is more difficult and is often not done with the necessary care (see e.g. Norde 2012). The main advantage of using corpora in this respect has been to look at these changes in larger sets of data. The possibility to look at authentic contexts has prompted a better understanding of the semantic-pragmatic mechanisms of change. The changes in form that we look for have not changed. More advanced quantitative corpus studies tend not to deal with progressive grammaticalization in individual items, but with competition between forms such as Nesselhauf (2006) and Tagliamonte et al. (2014) on grammaticalization posterchild *be going to* as future auxiliary, in competition with other auxiliaries. However, it is not just ease of identification that is the heart of the matter: working with a set of criteria that doesn’t cover the full range of items falling under the semantic definition of grammaticalization means that we are only able to confidently identify a subset of instances of grammaticalization, and that generalization and theory formation on the mechanisms of semantic change is based on this subset.

### Aim of this workshop

What is the way to tackle this issue and to develop a more extended toolkit of empirical changes to help the identification of semantic change? The papers on the subjectification of the Spanish verb *salirse* and the construction *a pesar de que* by Aaron and Torres Cacoullós (2005) and Torres Cacoullós and Schwenter (2006) show us the way to go. The aim of this workshop is to collect similar micro-analytical case studies for semantic change in English, with the hope of arriving at a critical mass allowing generalization of individual distributional evidence. Recent empirical studies of semantic change in the English noun phrase can be taken as one starting point. The noun phrase is mostly organized in an analytical way (only number and possession being marked morphologically) and hence an area in which it is difficult to provide empirical evidence for semantic change, using for example Lehmann's criteria for grammaticalization. However, a range of studies (Adamson 2000, Breban 2010) have shown that change in function (including grammaticalization and subjectification) goes together with a change in an item's position in relation to other items in the noun phrase. Change in the collocational range of nouns that an item co-occurs with is another way to detect starting and ongoing semantic change (Paradis 2000, Vandewinkel and Davidse 2008, Ghesquière 2014). Vartiainen (2013) opens a new window on subjectification by showing that subjectified adjectives co-occur more with indefinite determination. These are changes at micro-level, less obvious to see in corpus data than e.g. fusion or reduction of word forms. However, they are often the only observable reflections of semantic change, and are being applied as tools to identify and provide evidence for function change in current work on the noun phrase. Questions that feed into the search for distributional evidence is often how to operationalize processes that have been associated with semantic change in theoretical papers, such as the operationalization of collocational expansion (Himmelman 2004) by Vandewinkel and Davidse (2008) and by Hilpert (2008) for the development of future auxiliaries in the verb phrase. Hilpert (2008) and Van Bergen (2013) on the grammaticalization of *uton* as an adhortative shows how a similar micro-analytical approach can be used in the verb phrase. At sentence level, Walkden (2013) might provide inspiration: he shows how the position of the verb provides evidence that a functional misinterpretation of Old English *hwæt* has pervaded earlier philological work as well as dictionary definitions.

The aim of this workshop is in the first place to bring together research following a similar micro-analytical approach to semantic change, especially in other areas of English grammar than the noun phrase. The goal is to build up a toolkit of form/distribution and meaning/function change associations that can be applied in the empirical study of semantic change. In addition to this empirical goal, the workshop also invites papers that show how a detailed analysis of form and distributional changes can improve our understanding of the workings and mechanisms of semantic change, as for example evidenced in the work of De Smet (2012) on diffusional change, or can be applied in quantitative studies (e.g. Hilpert 2008). It invites papers that discuss how a wider range of form/distribution changes can be used to further develop a Construction Grammar model of semantic change, in which function and form are separate but linked poles that define constructions and the changes they undergo (Hilpert 2013, Traugott and Trousdale 2013).

### Themes of this workshop

This workshop invites papers that report on

- case studies identifying form and distributional changes that accompany semantic change in all areas of English
- the operationalization of theoretical notions such as collocational expansion (Himmelman 2004), decategorialization (Hopper 1991), etc.

- case studies that explore how theoretical proposals for operationalization such as Boye and Harder (2012) for grammaticalization and Visconti (2012) for subjectification can be applied to corpus data
- case studies applying quantitative analysis to micro-analytical changes in form/distribution
- the implications of micro-analytical case studies for our understanding of the workings and mechanisms of semantic change
- the implications of form/distributional changes for our understanding of the mechanisms of semantic change, and/or a Construction Grammar model of change

### Bibliography

- Aaron, Jessi E. and Rena Torres Cacoulios. 2005. Quantitative measures of subjectification: a variationist study of Spanish *salir(se)*. *Cognitive Linguistics* 16: 607–633.
- Adamson, Sylvia. 2000. A lovely little example: word order options and category shift in the premodifying string. In Olga Fischer, Annette Rosenbach and Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 39–66. Amsterdam: John Benjamins.
- Boye, Kasper and Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88: 1–44.
- Bergen, Linda van. 2013. Let's talk about *uton*. In Andreas H. Jucker, Daniela Landert, Annina Seiler and Nicole N Studer-Joho (eds.), *Meaning in the History of English: Words and Texts in Context*, 1–27. Amsterdam/Philadelphia: John Benjamins.
- Breban, Tine. 2008. The grammaticalization and subjectification of English adjectives expressing difference into plurality/distributivity markers and quantifiers. *Folia Linguistica* 42: 259–306.
- Breban, Tine, Jeroen Vanbiesen, Kristin Davidse, Lieselotte Brems and Tanja Mortelmans. 2012. Introduction: New reflections on sources, outcomes, defining features and motivations of grammaticalization. In Davidse, Kristin, Tine Breban, Lieselotte Brems and Tanja Mortelmans (eds.), *Grammaticalization and Language Change: New Reflections*, 1–35. Amsterdam/Philadelphia: John Benjamins.
- Brems, Lieselotte, Lobke Ghesquière and Freek Van de Velde. 2014. *Intersubjectivity and intersubjectification in Grammar and Discourse: Theoretical and descriptive advances*. Amsterdam/ Philadelphia: John Benjamins
- De Smet, Hendrik and Jean-Christophe Verstraete. 2006. Coming to terms with subjectivity. *Cognitive Linguistics* 17: 365–392.
- De Smet, Hendrik. 2012. *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: Oxford University Press.
- Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 103–120. Amsterdam/Philadelphia: John Benjamins.
- Ghesquière, Lobke. 2014. *The Directionality of (Inter)subjectification in the English Noun Phrase: Pathways of Change*. Berlin/New York: De Gruyter Mouton.
- Hilpert, Martin. 2008. *Germanic Future Constructions: A Usage-Based Approach to Language Change*. Amsterdam/Philadelphia: John Benjamins.
- Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press
- Himmelmann, Nikolaus. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In Walter Bisang, Nikolaus Himmelmann and Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 21–42. Berlin: Mouton de Gruyter.

- Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth C. Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization. Volume I: Focus on Theoretical and Methodological Issues*, 17–35. Amsterdam: Benjamins.
- Lehmann, Christian. 1995 [1982]. *Thoughts on Grammaticalization*. München: LINCOM Europe. [First published as *Thoughts on Grammaticalization: A Programmatic Sketch*. (Arbeiten des Kölner Universalien-Projektes.) University of Cologne, Institut für Sprachwissenschaft.]
- López-Couso, María José. 2010. Subjectification and intersubjectification. In Andreas H. Jucker and Irma Taavitsainen (eds.), *Handbook of Historical Pragmatics*, 127–163. Berlin: Mouton de Gruyter.
- Nesselhauf, Nadja. 2006. The decline of *be to* and the rise of *be going to* in Late Modern English: connection or coincidence? In Christoph Houswitschka, Gabriele Knappe and Anja Müller (eds.), *Anglistentag 2005 Bamberg. Proceedings*, 515–529. Trier: WVT.
- Norde, Muriel. 2012. Lehmann's parameters revisited. In Kristin Davidse, Tine Breban, Lieselotte Brems and Tanja Mortelmans (eds.), *Grammaticalization and Language Change: New Reflections*, 73–110. Amsterdam: Benjamins.
- Paradis, Carita. 2000. Reinforcing adjectives. A cognitive semantic perspective on grammaticalization. In Ricardo Bermúdez-Otero, David Denison, Richard M. Hogg and C. B. McCully (eds.), *Generative theory and corpus studies*, 233–258. New York: Mouton de Gruyter.
- Tagliamonte, Sali A., Mercedes Durham and Jennifer Smith. 2014. Grammaticalization at an early stage: Future *be going to* in conservative British dialects. *English Language and Linguistics* 18: 75–108.
- Torres Cacoullou, Rena and Scott Schwenter. 2006. Towards an operational notion of subjectification. In Rebecca T. Cover and Yuni Kim (eds.), *Proceedings of the 31<sup>st</sup> Annual Meeting of the Berkeley Linguistics Society*, 347–358. Berkeley: Berkeley Linguistics Society.
- Traugott, Elizabeth C. 2010. (Inter)subjectivity and (inter)subjectification: a reassessment. In Kristin Davidse, Lieven Vandelanotte and Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 29–69. Berlin: Mouton de Gruyter.
- Traugott, Elizabeth C. 2012. The status of onset contexts in analysis of micro-changes. In Merja Kytö (ed.), *English Corpus Linguistics: Crossing Paths*, 221–255. Amsterdam: Rodopi.
- Traugott, Elizabeth C. and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Vandewinkel, Sigi and Kristin Davidse. 2008. The interlocking paths of development towards emphatic adjective *pure*. *Journal of Historical Pragmatics* 9: 255–287.
- Vartiainen, Turo. 2013. Subjectivity, indefiniteness and semantic change. *English Language and Linguistics* 17: 157–179.
- Visconti, Jacqueline. 2013. Facets of subjectification. *Language Sciences* 36: 7–17.
- Walkden, George. 2013. The status of *hwæt* in Old English. *English Language and Linguistics* 17: 465–488.