

On the subordinating conjunctions as discourse markers in Lithuanian

Erika Jasionytė-Mikučionienė
(University of Vilnius)

Keywords: [pragmaticalization](#), [multifunctionality](#)

Schedule: [Fri 14.30 Room 9](#)

Over the last decades the development of conjunctions into discourse markers has been researched by a number of scholars (Degand & Simon-Vandenberg 2011, Rawoens 2015, among others). In Lithuanian, the studies on discourse markers are fragmented and lack the application of efficient qualitative and quantitative methods of analysis. This paper focuses on the Lithuanian discourse markers *kad* ‘that’ and *net* ‘until, even’ as well as their combinations with other particles and/or conjunctions (such as *bet (tai) kad; na/nu (tai) kad; tai kad; kad ir; net ir* etc.), which have not received an in-depth analysis so far, except for some observations made in Pajėdienė (2010), Sawicki (2012), Holvoet (2015) and Valančė (2017). Adopting both a synchronic and diachronic perspective and applying corpus-driven methodology, the present study aims at investigating the semantic functional potential of *kad* and *net* as well as their position and structural status in discourse. The data set analyzed includes both written (i. e. fiction) as well as spoken Contemporary Lithuanian and Old Lithuanian texts (16th–17th centuries).

The results of the analysis are interpreted within the framework of pragmaticalization (cf. Diewald 2011). It is shown that *kad* and *net* undergo semantic bleaching, (inter)subjectification and layering. In the earliest Lithuanian texts, *kad* is most often used in different types of adverbial clauses, including time, purpose and conditional clauses; *net*, in its turn, functions as a temporal conjunction or particle. However, in Contemporary Lithuanian, *kad* and *net* are multifunctional: *kad* can be used either as a conjunction or, in certain contexts (especially in dialogues and sentence initial position) as a discourse marker with a clear discourse function – to signal a discourse shift and to preface a response or reaction of the speaker, cf.:

A. *Tai gal greitai grįžti žadėjo?*

B. ***Kad nežinau, – gūžtelėjo pečiais Silvija.*** (CCLL-Fic)

‘A. So maybe he promised to return soon? B. Well, I don’t know, – Silvia shrugged her shoulders.’

As has been observed by Sawicki (2012, 151), the utterances opening with *kad* typically convey “various shades of scepticism, reservations or uneasiness of the speaker about the content of the previous turn and offer justifications or excuses for not complying or obeying”. However, the analysed data show that *kad*, besides its attitudinal functions, can also be associated with additional pragmatic functions, including emphasis and exemplification, i. e. addition of more specific information to the previous utterance. *Net*, as a discourse marker, is positionally mobile (cf. Simon-Vandenberg & Willems 2011, 361) and primarily marks emphasis or speaker’s surprise, cf.:

– *[K]aip tu čia šneki, sakau. Negaliu klausyti iš viso, taip kažkaip net.* (CCLL-Sp)

‘– What are you saying now, I say. I cannot listen to this at all, actually’

The data support the hypothesis made by Aijmer that “pragmatic markers have meaning potentials rather than fixed meanings which are realized in the same way in all situations” (Rühlemann & Aijmer 2015, 18), i. e. they have core meanings that can be modified in interactive contexts.

References:

- Degand, Liesbeth & Anne-Marie Simon-Vandenberg 2011, Introduction: Grammaticalization and (inter)subjectification of discourse markers, *Linguistics* 49(2), 287–294.
- Diewald, Gabriele 2011, Pragmaticalization (defined) as grammaticalization of discourse functions, *Linguistics* 49(2), 365–390.
- Holvoet, Axel 2015, Semantic functions of complementizers in Baltic, K. Boye, P. Kehayov (eds.), *Complementizer Semantics in European Languages*, Berlin/Boston: Mouton de Gruyter, 225–263.
- Pajėdienė, Jūratė 2010, Prijungiamieji laiko ribos sakiniai su jungiamuoju žodžiu *net* senojoje lietuvių kalboje, *Baltistica XLV(2)*, 265–283.
- Rawoens, Gudrun 2015, The Swedish connective *så att* ‘so that’: From subordinator to discourse marker, A. D. M. Smith, G. Trousdale, R. Walthereit (eds.), *New Directions in Grammaticalization Research*, Amsterdam/Philadelphia: John Benjamins, 51–65.
- Rühlemann, Christoph & Karin Aijmer 2015, Introduction. Corpus pragmatics: laying the foundations, K. Aijmer, Ch. Rühlemann (eds.), *Corpus Pragmatics: A Handbook*, Cambridge: Cambridge University Press, 1–26.
- Sawicki, Lea 2012, Responsive discourse particles in Lithuanian dialog, *Baltic Linguistics* 3, 151–175.
- Valančė, Danguolė 2017, *Pabrėžiamųjų dalelyčių funkcijos ir vartoseną dabartinėje rašytinėje lietuvių kalboje*, PhD thesis, Kaunas: Vytautas Magnus University.
- Simon-Vandenberg, Anne-Marie & Dominique Willems 2011, Cross-linguistic data as evidence in the grammaticalization debate: The case of discourse markers, *Linguistics* 49(2), 333–364.



Continuous vector space models for variation and change in sparse, richly annotated Indo-European argument structure data

Gard B. Jensen, Jóhanna Barðdal, Laura Bruno, Esther Le Mair, Peter Alexander Kerkhof, Svetlana Kleyner, Leonid Kulikov & Roland Pooth

(Ghent University; Ghent University, Oxford Research Centre in the Humanities; Ghent University; Ghent University; Ghent University; Ghent University & Ghent University)

Keywords: case variation, computational linguistics, historical–comparative linguistics

Schedule: Fri 16.30 Room 4

Quantitative studies of variation and change for historical languages are often hampered by sparsity of attested data but with rich annotation drawing on long traditions of linguistic and philological scholarship (Jensen & McGillivray 2017). Conversely, in natural language processing (NLP) for modern languages, data are plentiful but annotated data are scarce, prompting the use of neural network models that can accurately infer linguistic properties based on distributional information from very large un-annotated corpora (Mikolov et al. 2013a, Mikolov et al. 2013b). These techniques are relevant to historical linguistics because of their ability to handle sparse data and to model highly complex relations. Distributional approaches to variation and change in historical data (Barðdal et al. 2012, Jensen 2013) have previously relied on vector space representations that capture broad patterns but may struggle with highly complex distributional relations with sparse data.