

Secondary Publication



Birzer, Sandra

7 Competing constructions and language contact : Slavic discourse-structuring elements on the basis of non-finite verba dicendi

Date of secondary publication: 09.12.2024

Submitted Version (Preprint), Bookpart

Persistent identifier: urn:nbn:de:bvb:473-irb-1052658

Primary publication

Birzer, Sandra (2024): 7 Competing constructions and language contact : Slavic discourse-structuring elements on the basis of non-finite verba dicendi, in: Elena Graf und Ulrich Schweier (Hrsg.), Pragmaticalization : Language Change between Text and Grammar, Berlin/Boston: De Gruyter, pp. 229–282, doi: 10.1515/9783110761139-008.

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available with all rights reserved.

7. Competing constructions and language contact: Slavic discourse-structuring elements on the basis of non-finite *verba dicendi*

Abstract: Based on corpus data from Croatian, Czech, Polish and Russian, this paper explores the role of language contact in the historical development of “discourse structuring elements” (DSEs) based on non-finite *verba dicendi* such as Croatian *iskreno.ADV govoreći.AP* ‘frankly speaking’ or Czech *upřímně.ADV řečeno.PTCP.SG.N* ‘honestly speaking (lit. spoken)’ and assesses the development with regard to grammaticalization, lexicalization and pragmaticalization. All four languages feature two non-finite verb forms as formants. One of them is the adverbial participle, which functions as a formant in all four object languages and may thus be considered the originary Slavic construction. The second verbal formant, namely the past participle passive in Croatian and Czech (e.g. Croatian *iskreno.ADV řečeno* or Czech *upřímně.ADV řečeno* ‘frankly speaking’) and the infinitive in Russian (e.g. *tak skazat’* ‘so to say’) has arisen via pattern replication from the source languages German and French respectively. Polish, in contrast, features the perfective adverbial participle as a second formant, but in earlier stages used matter replication from Latin (e.g. *generalnie.ADV* vs. *ogólnie.ADV mowiac* ‘generally speaking’) to fill the adverbial slot of the construction.

In all object languages, only one of the two verbal formants is part of a semi-productive construction. Which construction becomes semi-productive depends on language-internal factors.

On the synchronic level, the DSEs fulfill four domains of function:

- a) pragmatic (e.g. Russian *čestno govorja* ‘frankly speaking’);
- b) contextualization (e.g. Czech *řečeno jazykem teologie* ‘speaking (lit. said) in the language of theology’);
- c) quotative indexing with the prototypical construction *govoreći / řečeno* + noun.ins+ noun.gen (e.g. Croatian *govoreći rječnikom sv. Pavla* ‘lit. speaking in the language of St. Paul’);
- d) direct speech marking with the help of “plain” *govoreći*, which only occurs in Croatian.

The DSEs with contextualising or quotative functions predominantly show a behaviour that is indicative of a grammaticalization process, whereas the DSEs from the pragmatic domain display a behaviour typical of lexicalization with regard to integrity and paradigmaticity. Their paradigmatic and syntagmatic variability and the leap-like increase of DSEs in the 18th and 19th centuries neatly fits in with co-optation, but rather not with grammaticalization (which Heine 2013 claims to necessarily follow co-optation). Since the development of the DSEs based on non-finite *verba dicendi* as an entire group cannot be encompassed by the analytical apparatus of grammaticalization, lexicalization and pragmaticalization research so far, we propose to apply Beijering’s (2012) model of composite change. This model allows to form four clusters of DSEs, whose features are characteristic for grammaticalization, lexicalization, pragmaticalization and co-optation respectively, with the latter forming the interface between the clusters.

Keywords: discourse-structuring elements, Croatian language, Czech language, Polish language, Russian language, grammaticalization, lexicalization, pragmaticalization, co-optation, corpus-based study

0 Introduction

The study of discourse-structuring elements (DSEs) – known in the respective individual Slavic philologies under such different terms as Polish *komentarz metatekstowy* ‘metatextual comment’, Croatian *diskursivan marker* ‘discourse marker’, Czech *diskurzni marker* ‘discourse marker’ or *diskurzni částica* ‘discourse particle’ or Russian *modal’naja částica* ‘modal particle’ – has a long tradition in Slavic studies. This contribution focuses on a specific subtype of discourse-structuring elements, namely those with a non-finite *verbum dicendi* as basis, and addresses the role of language contact in their historical development. To this end, we will relate Czech data from the historical subcorpora of the Czech National Corpus (henceforth CzNC) to the results of previously conducted studies on Russian (Birzer 2012, 2013), Polish (Birzer 2017a, b) and Croatian (Birzer 2018a).

This paper is structured as follows: Section 1. provides a cross-linguistic survey of the distinctive features of DSEs, whereas section 2. addresses methodological issues. Section 3. explains how the DSEs for this study were identified in the individual object languages and

what con- and divergencies they display. On this basis, we will draw conclusions on the role of language contact in their historical development (section 4.) and finally, we will discuss whether the emergence of this specific type of DSEs can be characterised as a process of grammaticalization, lexicalization or pragmaticalization (section 5.).

1 Discourse-structuring elements on the basis of non-finite *verba dicendi*

Metatekst and *discourse* are probably the two buzz words that mark the beginning of this branch of research. At the same time, they stand symbolically for the orientation of Eastern (cf. the seminal publication Wierzbicka 1971) and Western linguistic schools (cf. Schiffrin 1987). To make things worse, a multitude of terms circulate for the object of study – e.g. *discourse marker*, *discourse particle*, Polish *komentarz metatekstowy* ‘metatextual comment’, Croatian *diskursivan marker* ‘discourse marker’, Czech *diskurzni částica* ‘discourse particle’ or Russian *modal'naja částica* ‘modal particle’, to name but a few. As these terms are sometimes ambiguous and do not offer any coherent defining criteria,¹ delimiting the object of study is a challenge. Consequently, we have chosen to refer to our object of study as discourse structuring elements (DSEs) and to develop a catalogue of criteria for identifying DSEs. We have opted for this term as *element* neither implies membership in a particular word class nor a specific syntagmatic size; at the same time, *discourse-structuring* covers the following usage contexts (cf. the case study Birzer 2015a):

- the marking of reformulations within the narration (cf. del Saz and Fraser 2003),
- the distinction between actual narration and the speaker’s attitude to events in the narration, introduced by a reformulation marker,
- the distinction between narration and quotations embedded in it,
- marking the resumption of narrative strands or the introduction of digressions.

All the units termed so differently (for the different terms and the underlying line of argumentation cf. e.g. Fraser 1996, 1999; Fischer 2006; Grochowski 1986a, 1986b, 1998; Grochowski, Kisiel and Żabowska 2014 to mention just a few) occur in these contexts. The term *DSE* is thus to be understood as a neutral term, which is not restricted to our object of investigation in the narrower sense, i.e. DSEs based on non-finite *verba dicendi*, but also includes the units previously referred to as *discourse markers*, *pragmatic markers* or *discourse particles* (and their translation equivalents in various Slavic languages).

The characteristic features for identifying DSEs are graded into a) general features that apply to all kinds of DSEs and b) features that are specific to DSEs based on non-finite *verba dicendi*. The latter derive from the (morpho)syntactic specificities of this subgroup. The general features largely correspond to the features described by Rathmayr (1985) for particles in the function of pragmaxemes. In their discussion we will also draw parallels to approaches from general linguistics or to linguistic research on other, non-Slavic object languages.

General features of DSEs are:

a) syntactic integration into an utterance (host utterance).

This criterion is important to set off DSEs from clauses or sentences with discourse structuring function: Fischer (2006: 8–9) points out that even whole sentences can have

¹ As an example of ambiguity, consider the statement in Grochowski, Kisiel and Żabowska (2014: 26–27): “Należy jeszcze raz podkreślić, że partykulami w przyjętym tu rozumieniu nie są tym samym wyrazy i morfemy figurujące na listach “partykuł” w gramatykach szkolnych” [It should be stressed once again that particles in the sense applied here do not equal to the words and morphemes included in the lists of “particles” in school grammars – translation S.B.]. Since the various terms used in the respective national philologies and their implications are discussed in Birzer (2012) for Russian, Birzer (2018a) for Croatian and Birzer (2017a) for Polish, we will refrain from discussing them here.

discourse structuring functions and thus the total set of all discourse structuring constructions is theoretically infinite. Fraser (2006: 189) also mentions syntactic integration as a fundamental characteristic. Two features proposed by Rathmayr (1985), namely irrelevance to the proposition and the lacking possibility to function as answer of focus questions, imply that Rathmayr also assumes syntactic integration as general feature of DSEs.

However, it must be pointed out that syntactic integration is only a relative one, as can be seen in the positioning of enclitics: In Croatian, enclitics display the so-called second position placement, i.e. they follow after the first syntagma. If the syntagma consists of several (stressed) elements, an enclitic may also split it and the enclitic will follow after the first stressed word. The latter is recommended in the Croatian standard; however, as Reinkowski (2001) shows, the implementation is subject to variation. Exceptions also apply to clause-initial phrases, since enclitics cannot be attached to elements that are separated by an intonation pause – represented by a comma on the orthographic level – from the rest of the clause (Reinkowski 2008: 132). Likewise, enclitics cannot be placed after long initial phrases ('heavy constituent constraint' according to Radanović-Kocić 1996: 435). Since this syntactic behaviour is relevant for DSEs in general and applies to our subgroup in particular, one may expect that DSEs do not host enclitics, which in turn can be taken as an indication for weak syntactic integration. The data from the HNK_30 corpus (http://filip.ffzg.hr/cgi-bin/run.cgi/first_form) support this assumption: The clitic reflexive pronoun *se* does not occur after the DSE, but another phraseme (1). Similarly, only the full forms of personal pronouns occur after the DSE (2–3), while the clitic forms follow another phraseme (4); the same behaviour is found with the clitic forms of auxiliaries (3, 5):

(1) Croatian

Iskreno govoreći mi se bunimo protiv toga predsjednika ...
 frankly speaking we.NOM REFL rebel.PRS.1PL against this president
 'Frankly speaking, we are rebelling against this president.'
 (HNK_30 corpus, text fo2003br61cl1517)

(2) Croatian

On je tamo sasvim zadovoljio, ali iskreno govoreći, njega bismo više voljeli vidjeti u veznom redu, jer su mu kapaciteti mnogo veći.
 he was quite satisfied there, but frankly speaking him.FF
 we would prefer to see in the midfield because his capacities are much higher
 'He was quite satisfied there, but frankly speaking, we would prefer to see him in the
 midfield because his capacities are much higher.'
 (HNK_30 corpus, text gs20030502sp20088)

(3) Croatian

Iskreno rečeno, meni je na neki način drago da se ovdje stalno nešto događa - rekla nam je ta starija gospođa.
 honestly spoken me.FF It's dear in a way that something is happening all the time
 this elderly lady told us
 'Honestly, it's dear to me in a way that something is happening all the time – this elderly
 lady told us'
 (HNK_30 corpus, text gs20020711os12181)

(4) Croatian

Iskreno govoreći već me jako dugo neki roman nije tako uzбудio
 honestly speaking yet me.CF how long some novel has not excited so much

pa još prvijenac.
and it's still the first novel

- (5) Croatian
Iskreno rečeno, nadali smo se pobjedi, ...
honestly spoken hope.PTCP.PL AUX.1PL REFL.ACC.CLITIC victory.DAT
'To be honest, we hoped for a victory, ...'
(HNK_30 corpus, text vj20001023sp02).

Czech clitics, in contrast, behave differently from the Croatian ones, as they take the position immediately after the DSE, both in intrasentential (6) and sentence-initial position (7). Quite interestingly, our corpus research provided roughly ten times more matches of non-detached DSEs than DSEs detached from its host clause via punctuation (8–9).² One may cautiously interpret this as an indication that the Czech DSEs are possibly integrated more closely into their host clause than the DSEs in the remaining object languages:

- (6) Czech
A pak, když zatarasí cestu automobilu, mají možnost chytit buď amerického turistu,
and then , when they block the path of a car , they have the opportunity to catch either
amerického turistu, nebo bohatého jehož předkové, mírně řečeno,
Španěla,
an American tourist or a rich Spaniard whose ancestor.NOM.PL mildly spoken
se měli chovat líp.
REFL.ACC.CF have.PTCP.PL behave better
'And then, when they block the path of a car, they have the opportunity to catch either
an American tourist or a rich Spaniard whose ancestors, to put it mildly, should have
behaved better.'
(Bradbury, Ray. 1989. Kaleidoskop a jiné povídky.)
- (7) Czech
Přesně řečeno, mě ženy moc nepřitahují.
strictly spoken me.ACC.CF woman.NOM.PL very not_attract.PRS.3PL
'Strictly speaking, I'm not really attracted to women.'
(Murakami, Haruki. 2015. Bezbarvý Cukuru Tazaki a jeho léta putování).
- (8) Czech
Upřímně řečeno mi připadá, že je z tebe celá rozrušená.
honestly spoken me.DAT.CF seem.PRS.3SG Šshe's upset with you.
'Honestly, I think she's upset with you.'
(Cole, Martina. 2013. Bestie)

² Examples are taken from the syn2020 subcorpus of the Czech National Corpus with the help of the following queries:

[tag="D.*"] [word="řečeno"] [word="se"]

[tag="D.*"] [word="řečeno"] [word=","] [word="se"]

[tag="D.*"] [word="řečeno"] [word=","] [tag=".H.*"] 4 matches

[tag="D.*"] [word="řečeno"] [tag=".H.*"] 44 matches

[tag="D.*"] [word="se"] [word="řečeno"] and [tag="D.*"] [tag=".H.*"] [word="řečeno"] gave no matches

(9) Czech

<i>Upřímně řečeno</i>	<i>mě</i>	<i>to</i>	<i>fakt</i>	<i>vystrašilo.</i>
honestly spoken	me.ACC.CF	DET.NOM.SG.N	indeed	scare.PST.SG.N

‘Honestly, it really scared me.’
(LaBanová, Elizabeth. 2014. Esej o tragédii.).

Now, with regard to the criterion of phrasal length in particular, one could argue that it has no informative value with regard to syntactic integration, since phrasal length does not generally correlate with the degree of syntactic integration.

In this context, the data on the Polish enclitic conditional marker *by* are very valuable, since in Polish initial phrasemes do not show deviant behaviour and phraseme length also has no effects (cf. Hansen 2010: 345 on the positioning of the conditional marker *by*). In relation to DSEs, *by* shows the same behaviour (10) as Croatian clitics:³

(10) Polish

<i>Prawdę</i>	<i>mówiąc</i>	<i>nigdy</i>	<i>by-m</i>	<i>tam nie chodził ...</i>
truth.ACC	speak.AP	never	COND-1SG	not go there

‘Honestly speaking, I would never go there.’
(NKJP, R. Dudek. 1997. Mięśny książkę.).

We may deduce from the discussion that in all object languages, syntactic integration into the utterance can be observed for DSEs, but is actually weak. A possible explanation is provided by the two following features:

b) irrelevance to the proposition and, consequently, the possibility of syntactic elimination (cf. Rathmayr 1985: 42).

This feature is also mentioned by Fraser (2006: 189), which underlines its general validness beyond Slavic data.

c) Impossibility to function as an answer to a focus question (cf. Rathmayr 1985: 72).

Within the group of generally valid features of DSEs, features a)–c) once again form a subgroup of their own, since they are valid without restriction, while the following three features d)–f) are also generally valid, but individual groups of DSEs require their differentiated consideration.

d) Lack of negatability (cf. Rathmayr 1985: 72)

It is important to note that DSEs based on *verba dicendi* comprise several components, so in theory, different negation scopūs are possible – on the one hand, over the DSE as a whole and, on the other hand, over one of the DSE components. If, in addition to the *verbum dicendi*, at least one more component is autosemantic – as is the case with the constructions ADVERB + ADVERBIAL PARTICIPLE, ADVERB + PARTICIPLE E and ADVERB + INFINITIVE – a partial negation scopus is indeed possible (Polish *mówiąc nie delikatnie, a szczerze* ‘speaking not mildly, but frankly’ resp. Croatian *ne blago, a iskreno rečeno* ‘speaking (lit. said) not mildly, but frankly’ resp. Russian *ne mjagko, a čestno skazat* ‘speaking (lit. to speak) not mildly, but frankly’).⁴

e) DSEs cannot be under stress (cf. Rathmayr 1985: 72).

Most possibly, Rathmayr’s work on particles inspired her to develop this criterion: Particles usually have relatively little phonological substance, which favours the absence of stress (which in turn is a feature of cliticization; several particles, e.g. Russian *že* ‘yet’ have undergone cliticization). When it comes to DSEs based on *verba dicendi*, their mere phonological

³ Unless otherwise noted, all data on Polish come from the Polish National Corpus (NKJP, www.nkjp.pl).

⁴ For a detailed explanation see Birzer (2018, 2017b).

substance alone allows to presume the existence of secondary stress. This becomes even more apparent when we take into account the potential involvement of partial negation, which puts the focus and, consequently, stress on the DSE constituent introduced by a ‘but’, e.g. *mówiąc nie delikatnie, a szczerze* ‘speaking not mildly, but frankly’.

f) Accumulation with other DSEs (cf. Rathmayr 1985: 72).

Just like particles, DSEs may also accumulate. Especially noteworthy is the function of concretizer (Konkretizator) described by Mendoza (2009: 983), which secondary connectors, including DSEs, may assume in relation to primary, multifunctional conjunctions. This function invariably leads to the accumulation of DSEs (in connective function) and may also encompass conjunctions proper. Likewise, an accumulation of DSEs and particles is also possible, cf. (11).⁵

(11) Polish

Szczerze mówiąc to chyba nie ma sensu ...
honestly.ADV speak.AP this.NOM probably.PART doesn't make sense
‘To be honest (lit. honestly speaking), this probably doesn't make sense...’
(NKJP, Trybuna Śląska, 2002-05-21).

2 Methodology

To understand how the chosen methodology is motivated and how it shapes the research object (cf. section 3.), it is necessary to give a brief outline of the state of research.

The existing literature on DSEs can be roughly divided into three main streams. While Western literature is mainly interested in pragmatics (see, for example, Fraser 1996, 1999, 2006), linguistic schools from Eastern Europe, especially from Poland (cf., among others, Wierzbicka 1971; Grochowski 1983, 1986a) are rather concerned with aspects of text linguistics. Due to the proximity between pragmatics and semantics, it is hardly surprising that in the field of system linguistics attention is paid primarily to semantic issues (cf. e.g. Mosegaard-Hansen 2006; Nyan 2006). Morphosyntactic aspects, on the other hand, are either not addressed at all or are described in an inadequate manner (the analyses of *on the one hand – on the other hand* by Fraser (2006) and of *technically* by Itani (1995) are symptomatic).

Since this contribution pursues a panchronic approach and traces how parenthetical DSEs based on infinite *verba dicendi* developed, morphosyntactic and semantic aspects play a central role: Synchronically, the DSEs present themselves as (predominantly) multifunctional parentheses whose individual functions are closely linked to specific syntactic and semantic contexts. In addition, the explications for the individual functions are linked to one another by semantic bridges, so semantic maps (cf. Birzer 2015a) can be drawn. The syntactic development of DSEs is characterized by the loss of clause status, for which the loss of subject co-reference or control is symptomatic. Finally, historical language contact lead to replicated constructions in all the languages studied. However, with regard to morphosyntax, the resulting constructions differ depending on the specific contact situation.

We may state that a morphosyntactic description of DSEs is still lacking for both synchrony and diachrony. Their historical semantic and pragmatic development has hardly been addressed as well. In Slavic historical linguistics, only Nikolaeva (2008) on the development of particles deserves mention. Furthermore, linguistic typology has begun to study discourse-structuring parentheses in more detail only a short time ago (cf. Dehé and Kavalova 2007; Kaltenböck, Heine, and Kuteva 2011).

We opted for a corpus-based approach to analyse the DSEs based on non-finite *verba dicendi*. Using corpora is indispensable as the relevant contexts for describing the development

⁵ For details on Russian and Polish, see Birzer (2015a); on Croatian, see Birzer (2018).

of DSEs can hardly be found in dictionaries or grammars: On the morphosyntactic level, the loss of clause status is the main characteristic of non-finite *verba dicendi* developing towards DSEs. Thus, it is essential to find bridging contexts in which the morphosyntactic behaviour of the respective *verba dicendi* oscillates between that of a clausal predicate and the component of a DSE. However, instances of oscillating behaviour are hardly represented in grammars and dictionaries since grammars and dictionaries with their inherent normativity usually provide only prototypical examples.

The number of annotated diachronic corpora varies for our object languages, as does the coverage of centuries and the amount of tokens available for each century. For Russian and Czech we used the diachronic subcorpora of the respective National Corpora, for Polish we used the PolDi and the KorBa corpus and a plain text collection from Bibliotheka pisarzy staropolskich (<http://www.staropolska.pl/bibliografia/bps/index.html>). For Croatian we used the historical subcorpora of the Croatian National Corpus and the CroDi corpus. Plain texts can be used because the texts need to be searched for a very limited number of non-finite forms of the respective *verba dicendi* in order to trace the development of the DSEs based on them.

Consequently, corpus data form the main source of our database, but traditional, qualitative sources such as dictionaries or grammars are also consulted as supplementary resources.

3 Constructions studied in the object languages

In the following, we will discuss the properties of the constructions to be studied and their relevance for methodological and theoretical issues. Mainly data from Russian will be adduced as illustrative material; the data from Polish, Czech and Croatian display similar behaviour, so the findings on Russian can be generalised.

The first reason for choosing DSEs based on SAY or SPEAK is frequency, which in turn is linked to the productivity of the respective construction and the degree of its “constructiveness”. Thus, in Russian, in addition to the DSEs based on *skazat* ‘say.INF’ and *govorja* ‘speak.AP’, there are other DSEs with non-finite bases such as *zabegaja* (e.g. *zabegaja vpered* ‘jumping ahead [lit. running ahead]’), *rezjumiruja* ‘summarize.AP’ or *podvedja (itog)* ‘draw.AP (a conclusion)’, of which *rezjumiruja* can be considered a *verbum dicendi* in the broadest sense. A search query targeting these DSEs in the *osnovnoj korpus* (main subcorpus) of the Russian National Corpus yields an exorbitantly higher token frequency of DSEs based on SAY and SPEAK (Tab. 1).⁶

⁶ The constructions shown in the table were also searched in reverse linearity of the components. The query for *rezjumiruja* ‘summarize.AP’ deliberately contained no object, since a previous query had shown that the adverbial participle with object complement occurs only in the function of secondary predication. With regard to sentence position, the default setting was that the queried string of elements must be framed by a punctuation mark immediately preceding respectively following the string. This ensured the parenthetical status of the matches and excluded as much data noise as possible. Since this query provided only one match for *rezjumiruja*, another query was made with only one punctuation mark immediately preceding the adverbial participle. Because the queries were intended to give a first, rough picture of the data, we refrained from manual postprocessing as the query syntax ensured a low proportion of data noise. Since all queries were conducted in the same corpus, the conversion to wpm was dispensed with.

	Absolute token frequency
ADV + <i>skazat</i> ‘say.INF’	4389
ADV + <i>govorja</i> ‘speak.AP’	6550
<i>zabegaja</i> ‘run.AP’ + ADV	149
<i>podvedja</i> ‘draw.PFV.AP’ + ACC	17
<i>podvodja</i> ‘draw.IPF.AP’ + ACC	31
<i>rezjumiruja</i> ‘summarize.AP’	128

Table 1. Absolute token frequency in the *osnovnoj korpus* of the Russian National Corpus

Furthermore, if we consider the items represented by each variable (e.g. ADV), more than 40 adverbials form DSEs based on *skazat* ‘say.INF’ (cf. Birzer 2013) and over 100 adverbials form DSEs based on *govorja* ‘speak.AP’ (cf. Birzer accepted, where a similar ratio is also described for Polish, and Birzer 2018a for Croatian). In contrast, for *zabegaja* ‘run.AP’ only the collocation *zabegaja vpered* ‘jump.AP (lit. run.AP) ahead’, possibly modified by *neskol’ko* ‘somewhat’, *nemnogo* ‘a bit’, *sil’no* ‘strongly’ etc. and for *podvedja* ‘draw.PFV.AP’ the doublets *podvedja.PFV.AP* / *podvodja.IPF.AP* *itog.GEN.PL* / *itogi.ACC.PL* ‘drawing conclusions’ are attested. In the case of *rezjumiruja* ‘summarize.AP’, we are even not so much dealing with collocations but rather with a wide range of syntagms that co-occur with the adverbial participle (12–14). It is also noteworthy that *rezjumiruja* has obviously not (yet) abandoned co-reference with the first argument of the matrix sentence, although impersonal matrix sentences like (12) are also possible. Thus, the range of syntagms co-occurring with *rezjumiruja* cannot be explained by the fact that *rezjumiruja* is part of a discourse-structuring construction with high productivity, but by its status as a secondary predicate, whose arguments are realized in various syntactic forms. In contrast, the DSEs based on SPEAKING and SAYING form a (semi)productive subgroup of all DSEs, and for this very reason their analysis is much more promising:

- (12) Russian
Rezjumiruja izložennoe vyše, prichoditsja priznat’, čto problema stracha smerti u podroستkov i junošej poka ostaetsja ves’ma zaputannoj.
‘Summarising the above, it must be acknowledged that the problem of death in adolescents and young people is still very confusing.’
Summarising the above, it must be acknowledged that the problem of death in adolescents and young people is still very confusing
- (13) Russian
Rezjumiruja ne sliškom veselye promežutočnye itogi analizov izotopnogo sostava vody v kometach, avtory stat’i v Science rešilis’ sdelat’ ves’ma otvetstvennoe umozaključenie: ...
‘Summing up the rather unpleasant interim results of analyses of the isotopic composition of water in the comets, the authors of the article in Science dared to draw a highly responsible conclusion: ...’
(T. Oganjesjan. Tajna zemnych okeanov // “Ėkspert”, 2015)
- (14) Russian
Kak by rezjumiruja vse eti fakty, odin iz ispolnitelej zaključaet: ...

‘As if to summarise all these facts, one of the performers concludes: ...’
(Zapis’ LiveJournal (2004)).

Let us now also consider the loss of co-reference in the two other DSE types discussed above. For *podvodja/podvedja itog/i* ‘drawing conclusions’, the majority of instances features impersonal matrix clauses or matrix verbs with a non-canonical subject (15); there are also rather rare cases where co-reference is at least doubtful (16).⁷ We may interpret this as evidence that *podvodja/podvedja itog/i* is losing clause status. In the case of *zabegaja* ‘jump.AP (lit. run)’, co-reference with the first argument of the matrix clause exists no more (17), as is also the case with the DSEs based on SPEAKING and SAYING:

(15) Russian

Podvedja itogi, nužno skazat’, čto každoe vystuplenie soprovoždalos’ burej suetlivogo vostorga i optimističeskim gromom aplodishmentov, čto, estestvenno, svidetel’stvuet o polnom edinodušii prisutstvujuščich.

‘To summarise, each presentation was accompanied by a storm of frantic enthusiasm and optimistic thunder of applause, which naturally indicated complete unanimity among those present.’

(S. Dovlatov. Remeslo. Povest’ v dvuch častjach. Čast’ 1. Nevidimaja kniga (1976))

(16) Russian

... on i formuet èti mini-modeli, a područnye potom uveličivajut ix do razmerov gigantskix kamennyx orjasin. Kamorka ego tesnaja, pod masštab figurok. U steny, kak glavnyj naturščik, belet ogromnyj čerep mamonta s malen’kimi dyrami glaznic i čej-to pozvonok razmerom v umyval’nuju rakovinu. Èto XX vek, podvodja itogi, uprjamo i skrupulezno iščet svoju ideju, naxodit provaly, začiščaet nerovnosti gipsa skal’pelem.

‘So he shapes these mini-models, and his assistants then scale them up to the size of giant, clumsy stone figures. His closet is cramped, to the scale of the figures. By the wall, like his main artistic model, a huge mammoth skull with small holes in its eye sockets and some vertebra the size of a washbasin are gleaming white. It is the twentieth century, summing up, stubbornly and scrupulously searching for its idea, finding the gaps, mending the unevenness of the plaster with a scalpel.’

(A. Voznesenskij. Na virtual’nom vetru (1998))

(17) Russian

I, zabegaja vperëd – èto byl pervyj i poslednij incident s mestnym naseleniem za vsju našu poezdku.

‘And, jumping ahead – this was the first and last incident with locals on our entire trip.’

(Otčët o velosipednom pochode (2001)).

Thus, the loss of clause status is a criterion for demarcating core and periphery among the DSEs. The semi-productive constructions on the basis of SPEAKING and SAYING form the core.

If it comes to verb semantics, *rezjimiruja* ‘summarize.AP’ is used in its direct meaning, whereas *podvodja / podvedja* ‘draw.AP’ and *zabegaja* ‘jump.AP (lit. run)’ are instances of metaphorical use. When SPEAKING and SAYING occur as element of DSEs, their original

⁷ In example (16) it remains unclear whether the narrator draws the conclusion that the situation in the artist’s closet represents the state of affairs by the end of the 20th century – which would be an instance of lacking co-reference – or whether the (anthropomorphized) 20th century, looking back and taking stock, re-evaluates ideas etc.

semantic component ‘oral articulation’ erodes (for more details, see Birzer 2012a). Both metaphorical use and semantic erosion are indicative of a grammaticalization process. Thus, analysing verb bases exhibiting semantic erosion will also be instructive for grammaticalization theory.

Verb semantics is also indicative of the registers the respective DSEs are used in. The verb *rezjumbovat* ‘summarize’ itself as well as the DSE *rezjumbova* ‘summarize.AP’ are typically used in texts representing the concept of written form (“konzeptuelle Schriftlichkeit” according to Koch and Oesterreicher); the same also applies to the DSEs based on *zabegaja* and *podvedja / podvodja*. Yet the DSEs based on SPEAKING and SAYING are not restricted in their usage contexts. A restricted usage context also seems to correlate with a limited number of functions: while the former DSEs serve primarily topic management, most of the latter are polyfunctional (cf. the explanations below and Birzer 2015a).

Summing up the findings from the previous paragraphs, we may conclude that the DSEs based on SPEAKING and SAYING form the core of the DSEs based on non-finite *verba dicendi*, as they combine high usage frequency, (semi)productivity, polyfunctionality and the loss of clause status. From the perspective of Construction Grammar, the core DSEs are an interesting research object because of their (semi)productivity and the competition between two construction types (in Russian the two types feature the non-finite bases *skazat*.INF ‘say’ and *govorja*.AP ‘speak’ respectively), which will be explained in more detail below. In the light of grammaticalization theory, the observed loss of clause status as well as the semantic erosion are worthwhile investigating.

In addition to the theoretic considerations, the core DSEs also lend to a diachronic corpus study much better than the DSEs from the periphery: The verbs with the meaning SPEAK and SAY, beginning with Old Church Slavonic *glagolati* ‘speak’ and *rešti* ‘say’, are consistently attested on the historic stages of all object languages and can be traced even in an untagged corpus. Thus, they offer the possibility of an encompassing panchronic analysis.⁸

As mentioned above, we may observe two different morphosyntactic patterns for the DSEs at the synchronic level of each object language. The representatives of the two formation patterns compete with each other; one of the two patterns is semi-productive and has a higher number of representatives. Formations with the imperfective adverbial participle occur in all four object languages. In Russian, they represent the more productive pattern compared to formations with the infinitive (i.e. *kstati govorja*.AP vs. *kstati skazat*.INF ‘by the way (lit. by the way speaking / to say)’); in Polish, they are semi-productive in contrast to isolated lexicalised items featuring the perfective adverbial participle (e.g. *prawdę mowiąc*.AP.IPF ‘honestly said (lit. telling the truth)’ vs. *prawdę powiedział*.AP.PFV ‘honestly said (lit. having said the truth)’). Croatian DSEs based on the imperfective adverbial participle compete with those based on the past participle passive (e.g. *iskreno govoreći*.AP.IPF ‘honestly speaking’ vs. *iskreno rečeno*.PTC.PASS ‘honestly said’); the pattern with the passive participle is semi-productive. The situation in Czech resembles that in Croatian, but in a more pronounced form: the imperfective adverbial participle (called *transgressive* in Czech grammaticography) has been retained in a single lexicalized DSE, namely *tak říkajíc*.AP ‘so to say (lit. saying so)’, whereas the construction based on the past participle passive is the productive pattern for all other DSEs (e.g. *upřímně řečeno*.PTCP.PST.PASS ‘frankly speaking (lit. spoken)’ or *jinak řečeno*.PTCP.PST.PASS ‘in other words (lit. otherwise spoken)’, but also *tak řečeno*.PTCP.PST.PASS ‘so to say (lit. said so)’).⁹ The role of language contact in the development of this competitive situation will be discussed in section 4.

⁸ Whenever the term DSE will be used in the following, it will refer to the DSEs based on non-finite forms of SAY and SPEAK, unless indicated otherwise.

⁹ For our corpus analysis we used the syn2020 subcorpus of the Czech National Corpus. The query for identifying DSEs matching the pattern ADV + *říkajíc*.AP was [tag="D.*"] [word="říkajíc"]; the second, very simple query [word="říkajíc"] confirmed that *tak říkajíc* is the only DSE based on *říkajíc* and that *říkajíc* occurs almost

At the synchronic level, the functional areas of the DSEs can be roughly divided into connectivity, the expression of speaker's stance, topic management (e.g. the introduction of digressions) – which will be summarised below under the label pragmatic functions –, the discursive contextualization of (parts of) a statement, the marking / ascription of quotations (quotative function) and the use as a filler. Constructions with morphosyntactically different modifiers are prototypical for specific functional areas (Table 2).¹⁰

functional area(s)	prototypical construction
pragmatic functions connectivity (17) expression of speaker's stance (18) <i>topic management</i> (19) filler (20)	ADV + NON-FINITE VERBUM DICENDI
discursive contextualization of (parts of) a statement (21)	ADJ.INST + LANGUAGE.INST + NON-FINITE VERBUM DICENDI
marking / ascription of quotations (quotative function) (22)	LANGUAGE.INST + PROPER_NAME.GEN + NON-FINITE VERBUM DICENDI

Table 2. Assignment of prototypical construction patterns to the functional areas of DSEs

- (17) Croatian
... nacionalizacija, konfiskacija, eksproprijacija . . . ukratko rečeno – ozakonjena otimačina, [se] rastegnula na mnoge hektare najbolje zemlje ...
 "... nationalization, confiscation, expropriation ... in short – legalized seizure, has extended itself over many hectares of best land."
 (HNK_30, text gs20021023sb14745)
- (18) Polish
Malownicza zakopiańska zima przerodziła się w wiosenne roztopy. Szczерze mówiąc, wolę skrzypiący śnieg pod stopami niż błoto ...
 'The picturesque Zakopane winter has turned into a spring melt. To be honest, I prefer the creaking snow underfoot to the mud ...'
 (NKJP, My "Polish Diary" // Dziennik Polski, 2002-02-01)
- (19) Russian
Ja ... byla preprovoždена v vannuju komnatu. — Kran sil'no ne otkryvajte, — velela Anja, — zabryzgaete vse, i mylo akkuratno položite. Ja ogljadela vystroennye po ranžiru, ideal'no vytertye banočki i flakončiki, polotenca, strogo parallel'no visjaščie na suške, fen, pomeščennyj v special'nuju podstavku, i ponjala: sud'ba zanesla menja k klassičeskoj staroj deve. Voobščę govorja, damy, nikogda ne vychodivšie zamuž, ne imevšie ni detej, ni ljubovnikov, k opredelennomu vozrastu raspadaјutsja na neskol'ko grupp. Odni pylajut strašnoj nenasiv'ju ko vsemu živomu Drugie, naobrot, prevraščajutsja v čadoljubivych tetušek, vospityvajuščich plemjannikov. Tak vot, vtorye, kak pravilo, bodrye optimistki, naplevatel'ski otnosjaščiesja k domašnemu hozjajstvu. ... Zato pervye prevraščajut čistotu v fetiš.

exclusively as part of this DSE, i.e. the function of secondary predicate is defunct. DSEs with the pattern ADV + *rečeno*.PTCP.PST.PASS were searched with the query [tag="D.*"] [word="rečeno"]. The search was conducted in March 2023.

¹⁰ Here "prototypical" means that the given construction fulfills this function in the majority of or even all object languages, although it is not necessarily the construction used most frequently for expressing the given function in the respective object language.

‘I ... was escorted to the bathroom. – “Don’t open the tap too much”, Anja told me, “you’ll splash everything, and place the soap carefully”. I looked at the ranked, perfectly wiped pots and bottles, at towels, hanging strictly parallel on the dryer, a hair dryer, placed in a special stand, and realized: fate has brought me to a classic old maid. Generally speaking, ladies who have never married, had no children, no lovers, and are of a certain age fall into two groups. Some are blazing with a terrible hatred for all living things ... Others, on the contrary, turn into child-loving aunts raising their nephews. Well, the latter tend to be cheerful optimists who don’t give a damn about the household. ... But the former turn cleanliness into a fetish.’
(RNC, Dar’ja Doncova. *Dollary carja Goroča* (2004)).

- (20) Russian
*Nu tuut vspominaetsja legenda/ **vobščę grja** o šamane// éta legenda u kogo/ u évenkov po-moemu/ u kakich-to severnych narodnostej//*
‘Well, there’s a legend that comes to mind // generally speaking about a shaman// that legend is of whom / of the Evenki I think / of some northern peoples //’
(RNC, *Beseda ob oborotnjach // Živaja reč’ ural’skogo goroda*, 1989).
- (21) Polish
*Jako bardzo ważna persona, a więc **mówiąc dzisiejszym językiem** – VIP, odtąd wszędzie miał mu towarzyszyć ochroniarz.*
‘As a very important person and therefore, speaking in today’s language, a VIP, he was henceforth to be accompanied everywhere by a bodyguard.’
(NKJP, Janusz Rolicki. 2002. *Edward Gierek: życie i narodziny legendy*).
- (22) Czech
*Tento živočich byl „**problémový krokodýl**“, **řčeno jazykem úředníků**.*
‘This animal was a “problem crocodile”, in the language of the officials.’
(CzNC, TextID nati1909:2).

The DSEs with the functions of discursive contextualization and marking / ascription of quotations are monofunctional (in the sense that they feature no pragmatic functions), have a compositional semantic structure and feature specific prototypical constructions. In contrast, many DSEs of the type ADV + NON-FINITE VERBUM DICENDI combine several functions and are semantically rather opaque. Quite noteworthy, this dichotomy also continues when it comes to the functions of larger classes of DSEs: the function of quotation marking (and possibly also that of quotation contextualization) relates to quotativity, i.e. a grammatical category, while the other functions are pragmatic in nature. This raises the question whether the exponents of the respective functions result from different processes of language change, namely grammaticalization, lexicalization or pragmaticalization, or whether they represent a process that combines elements of grammaticalization and lexicalization (cf. Beijering 2012). This will be discussed in section 5.

The range of pragmatic functions a DSE can fulfill is determined by the meaning of its non-verbal element – prototypically an adverb. In each of the object languages the respective (semi)productive construction that forms DSEs with pragmatic functions has produced around 100 DSEs. In order to grasp how semantics and pragmatic functions correlate, semantic explications for all DSEs of the type ADV + NON-FINITE VERBUM DICENDI were created following the model of the Moscow Semantic School (this and the following cf. Birzer 2018b). Additionally, the non-verbal elements of the DSEs were grouped into semantic hyperclasses. In the literature on discourse and pragmatic markers, there is disagreement as to whether all pragmatic functions of a DSE can be derived from one semantic explication, or whether each

pragmatic function requires a semantic explication of its own (for a more detailed account of the problem, see Birzer 2015a). Since our explications only served to identify subgroups of DSEs with the same pragmatic function, we opted for one (schematised) explication per DSE. Given around 100 DSEs with pragmatic functions in each object language, even this was not feasible for all object languages. Since new pragmatic functions often arise in oral language use, we used Polish DSEs with pragmatic functions to exemplify our approach, since Polish is one of the object languages whose national corpora also features a sub-corpus of oral texts.

It turned out (cf. the online available explications in Birzer 2018c) that the non-verbal elements can be assigned to four semantic hyperclasses RELATION, EVALUATION, COMPARISON and STRUCTURE, with each semantic hyperclass providing a specific range of functions. Thus, DSEs of the hyperclass COMPARISON have connective function, while the representatives of the hyperclass STRUCTURE combine the functions of topic management and the expression of speaker's stance. The hyperclass EVALUATION has the functions of connectivity and expressing speaker's stance; in addition to these functions the hyperclass RELATION may also serve topic management (cf. Diagram 1).¹¹

Functional sphere	topic management		speaker's stance		
	introduction; excursus; introduction of additional information; resumption of a topic	reformulating connective	introduction evaluation of a situation	hedging	filler
Individual functions	introduction; excursus; introduction of additional information; resumption of a topic		introduction evaluation of a situation	hedging	filler
Semantics (P and Q are variables representing semantic arguments in the explication)	P and Q share features	P and Q relate to each other	evaluation Q relates to P	evaluation Q is not generally valid	∅
Semantic hyperclass	STRUCTURE	COMPARISON	STRUCTURE	EVALUATION	RELATION

Diagram 1. Relationship between the functions of pragmatic DSEs and the semantic superclasses of their non-verbal elements.

4 Language contact

As has been mentioned above, language contact is the most likely reason why two (competing) patterns for the formation of DSEs exist at the synchronic stage of all object languages. As a rule, one patterns is (semi)productive and has a higher number of representatives. As constructions with the imperfective adverbial participle occur in all object languages (albeit with varying functions and usage frequency), they can be considered to be of genuinely Slavic origin. In Russian, they represent the (semi)productive pattern compared to formations with the

¹¹ For details refer to the case study Birzer (2015a), in which a representative of the semantic hyperclass RELATION, namely Polish *ogólnie mówiąc* 'generally speaking' and its Russian counterpart *вообще говоря* 'generally speaking', were analyzed for their polyfunctionality.

infinitive; the latter offers a parallel to the contact language French, where DSEs are also formed on the basis of infinitives, e.g. *à vrai dire*.INF ‘honestly speaking’ (lit. ‘to say honestly’).

In Polish, the construction with the imperfective adverbial participle is also (semi)productive in contrast to the formation pattern *(a)by X powiedzieć*.INF ‘in order to say X’ and isolated lexicalised formations based on the perfective adverbial participle (e.g. *prawdę mówiąc*.AP.IPF ‘speaking honestly (lit. speaking the truth)’ vs. *prawdę powiedział*.AP.PFV ‘honestly said (lit. having told the truth)’). German, which has been in contact with Polish for centuries, features the construction *um nicht zu sagen X* ‘not to speak about X’ (lit. ‘in order not to say X’), which can also be traced back to a final clause construction exploiting the infinitive.

In Croatian, DSEs based on the imperfective adverbial participle compete with those based on the past participle passive (e.g. *iskreno govoreći*.AP.IPF ‘honestly speaking’ vs. *iskreno rečeno*.PTCP.PST.PASS ‘honestly speaking (lit. said)’); the latter pattern is the (semi)productive one. German is also an important contact language for Croatian. Quite noteworthy, the situation in the two languages is parallel: German also has DSEs based on the past participle passive (*ehrllich / kurz / genauer gesagt*.PTCP.PST.PASS ‘honestly / briefly / more precisely speaking (lit. said)’ etc.), which is the (semi)productive pattern for forming DSEs in German.¹²

A similar pattern can be observed with Czech and its contact partner German. The construction based on the past participle passive is the (semi)productive pattern in Czech (e.g. *upřímně řečeno*.PTCP.PST.PASS ‘frankly speaking (lit. spoken)’ or *jinak řečeno*.PTCP.PST.PASS ‘in other words (lit. otherwise spoken)’ and displays a structural parallel to its German counterpart. Czech is probably the Slavic language that has had the longest and most extensive contact with German (cf. among others Berger 2014). In the following, we will use Czech data to illustrate how the influence of language contact can be traced.

For reasons of space, we will restrict our literature survey to the role of language contact in the development of DSEs. Since DSEs are characterised by specific morphosyntactic properties and, in addition, the question of whether their development is a case of grammaticalization, pragmaticalization or lexicalization is to be investigated, we chose the replication model (originally designed by Heine and Kuteva 2005) from a multitude of language contact models (for an overview from a Slavicist perspective cf. Rabus 2013: 36–56). It explicitly addresses the relation of language contact phenomena to grammaticalization theory (Heine and Kuteva 2003). Sakel (2007) develops the model further by distinguishing pattern replication – the equivalent of Heine and Kuteva’s grammatical replication – from matter replication and by discussing possible interactions between both types of replication. Finally, using typological data, Matras (2007a) provides borrowing hierarchies for function word types, e.g. for conjunctions, and the markers of grammatical categories.

Studies from general linguistics show that language contact exerts significant influence on the inventory of discourse markers (cf. foremost Matras 1998, 2007b). For items with connective function – a property of many DSEs – Matras observes the following behaviour in language contact situations:

- a. Connective particles are prone to direct borrowing (i.e. replication of form), particularly from a dominant language [...].
- b. Connectivity structures that operate at the level of the entire organization of the complex clause are prone to replication of that organization pattern. Such convergence or pattern replication occurs under the influence of a dominant language, but also in linguistic areas (where relations of prestige and dominance may not be obvious). What is replicated here is primarily the rules of form-function mapping, the ordering of elements and their grammatical meaning. In other words, we are dealing here with the replication or fusion of

¹² Italian is also an important contact language for Croatian. In the present case, however, it is more likely that German was the decisive contact language, since the Croatian and German constructions are identical in function and morphosyntax: both are based on the participle passive, while their Italian equivalent, exemplified e.g. by *a dire*.INF *il vero* ‘actually (lit. to say the truth)’, is based on the infinitive.

constructions (in the sense of Construction Grammar; cf. Goldberg 1995; Croft 2001). (Matras 2007b: 68).

From a formal point of view, the DSEs under investigation undoubtedly belong to Matras's type b. However, it needs to be clarified whether the respective DSEs are the result of language contact and, if so, whether they actually constitute pattern replications. In general, historical pattern replications in Slavic have received greater attention rather recently. For example, Wiemer and Hansen (2012) offer a first overview of contact-induced grammaticalization in Slavonic. Berger (2009) on the absentive as a result of German-Czech language contact and Zumstein (2012) on the con- and divergences between Old Church Slavonic and Greek passive constructions show that replicated constructions may be object both to a narrowed range of use (Berger 2009) and a further differentiation of their functions (Zumstein 2012). The volume Mendoza and Birzer (2022a) presents instances of contact between the Slavic vernaculars and the biblical languages Greek, Latin and Old Church Slavonic.

Many language contact models have in common that language contact is assumed to take place in a face-to-face situation. This assumption also applies to the replication model and to Matras' studies of DSEs in language contact situations. This can be explained by the fact that these contact models stem from the realm of language typology. They are grounded in the synthesis of data collected from languages of particular typological interest, often encompassing less(er)-known languages with scant, if any, written records. Consequently, researchers rely on synchronous (oral) data to construct models. This weakness is already pointed out by Rabus, who notes that

Situationen denkbar und auch im hier interessierenden Bereich historisch attestiert [sind], in welchen die Bedingung der direkten Interaktion mehrsprachiger Sprecher auf Diskursebene nicht gegeben ist, aber trotzdem kontaktinduzierter Sprachwandel festzustellen ist. [...] Die Ursache hierfür ist der Kontakt der Zielsprache mit einer kulturell und soziolinguistisch dominanten Schriftsprache, auch im Rahmen von Übersetzungen.

[Situations can be imagined and have been historically attested in our area of interest [inner-Slavic language contact – S.B.] to which direct interaction of multilingual speakers at the discourse level does not apply as condition, but in which contact-induced language change can nevertheless be observed. [...] The reason lies in the contact of the target language with a culturally and sociolinguistically dominant written language, also in the context of translations. [translation – S. B.] (Rabus 2013: 66).

In the case of contact between the Slavic languages and representatives of other language families, this type of language contact is even more pronounced. In the following, we will use the term “literacy contact” proposed by Verkholtantsev (2008) and refined by Rabus (2013), which denotes “language contact caused not by oral interaction but by the transmission and reception as well as translation of written texts [...]” (Rabus 2013: 66). Mendoza and Birzer (2022b) is a first attempt to systematize the various settings and outcomes of literacy contact in Slavic in a coherent approach.

Our present study may be considered a fragment in the larger mosaic of (historical) literacy contact. Its contribution to theoretical questions is rather limited. As previously mentioned, our primary hypothesis revolves around whether the competitive dynamics between constructions for forming DSEs can be interpreted as a consequence of language contact.

Rabus (2013: 64) posits that language contact can also amplify “certain developments, certain internal linguistic predispositions in the target language.” We will explore this phenomenon, known as the catalyst effect, in our data: after the replication of DSE-specific constructions, which lead to competition between the originary Slavic and the replicated construction, the number of DSEs based on these constructions increases drastically around the

turn of the 19th century (Birzer 2012a, 2017b, 2017c). We hypothesize that this increase is an instance of the catalyst effect.

4.1 Historical development of DSEs and the role of language contact

There is a historical explanation for the competition of constructions on the synchronic stage in which language contact and extralinguistic factors play a decisive role. Before the results of language contact are compared across languages, we will give a brief outline of the historical development in Russian, Polish and Croatian; Czech data will be discussed in more detail, as it has not been considered in previous publications.

Already in Old Church Slavonic a construction consisting of two *verba dicendi* serves to mark rendered speech and quotations. One *verbum dicendi* is semantically complex and denotes the manner of articulation (e.g. *moliti sja* ‘to pray’), while the second one is semantically less complex (*glagolati* ‘to speak’ and *rešti* ‘to say’, ‘to speak’) and marks the beginning of the speech rendition or the quotation. The second *verbum dicendi* may bear the form of participle (to be reanalysed as adverbial participle much later in the individual Slavic languages) or aorist (mainly of *rešti* ‘say’). One may assume that the influence of Greek has fostered the establishment of this construction, since syntactic and morphological similarities can be found in Old Church Slavonic and Greek parallel texts (cf. Birzer 2012b).

(23)	a.	Greek				
		καὶ	ἀνοιξας	τὸ	στόμα	αὐτοῦ
		CONJ	open.PTCP.AOR.NOM.SG.M	DET.ACC	mouth.ACC	his
		ἐδίδασκεν	αὐτοὺς	λέγων,		
		teach.IMPERF.3SG	them	say.PTCP.PRS.NOM.SG.M		
	b.	Old Church Slavonic				
		<i>i</i>	<i>otvr-z-</i>	<i>ousta</i>	<i>svoě</i>	
		CONJ	open.PTCP.PST.NOM.SG.M	mouth.ACC	REFL.POSS.ACC	
		<i>oučaaše</i>	<i>ja</i>	<i>glę: ...</i>	(Codex Zographensis)	
		teach.IMPERF.3SG	them	say.PTCP.PRS.NOM.SG.M		
		‘And he opened his mouth, and taught them, saying: ... ’ (Matthew 5:2; King James version)				

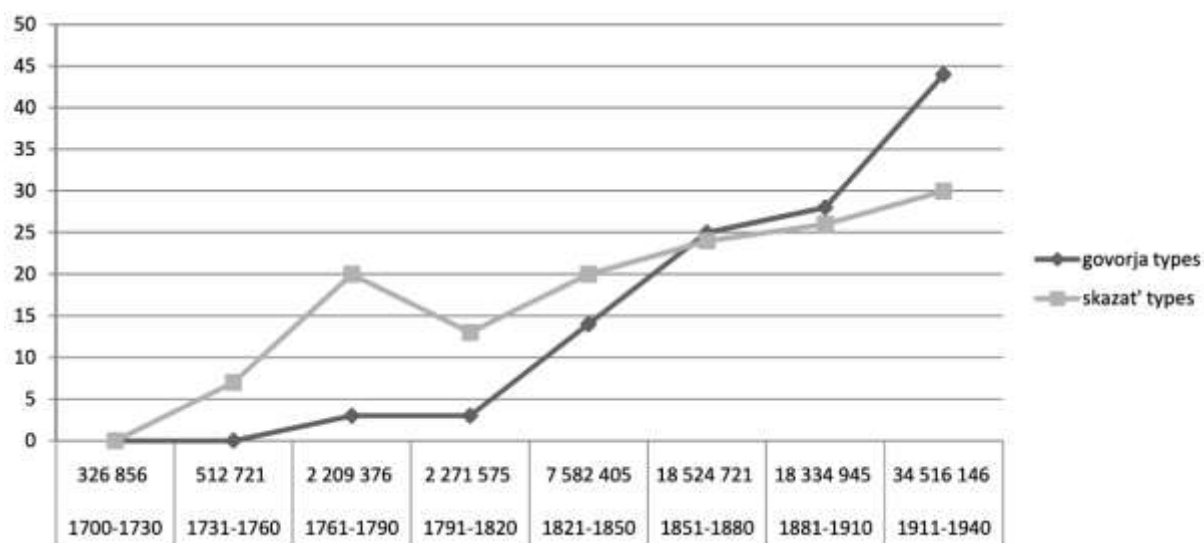
4.1.1 Russian

In Common East Slavic and Middle Russian the situation is quite similar to that in Old Church Slavonic, but first signs of the further development can be observed in Middle Russian: Deictic adverbs are attested as modification of the second *verbum dicendi* and in some cases, the second *verbum dicendi* “detaches” from its partner and moves into the clause or sentence which it marks as rendered speech or quotation, i.e. if it constitutes a non-finite verb form its status as secondary predicate is lost (for this and the following cf. Birzer 2012a):

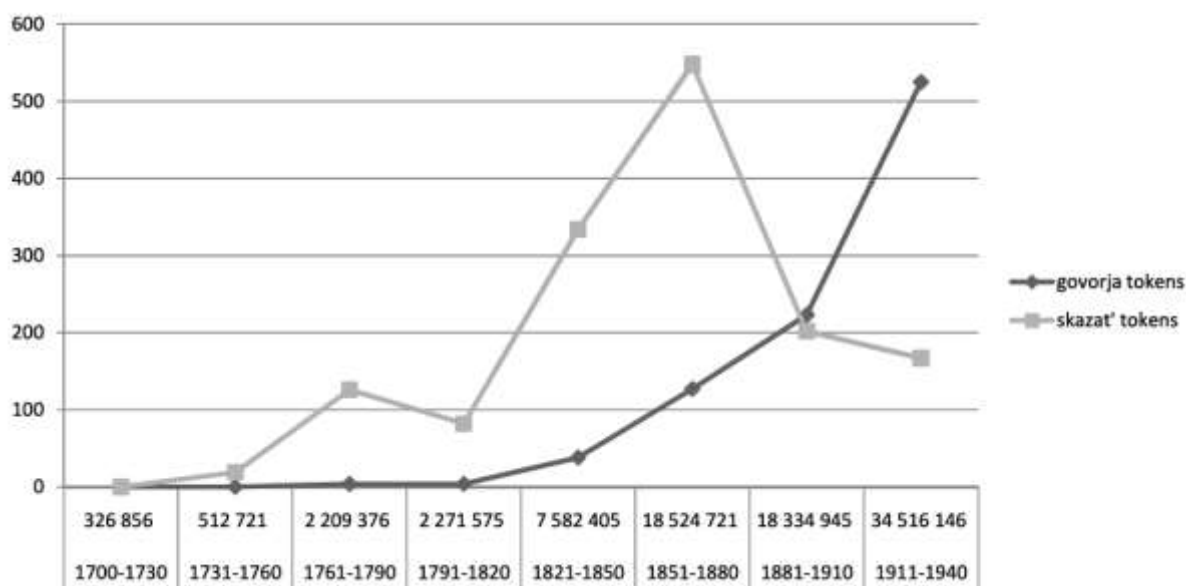
(24)	Common East Slavic [assumptions about when the end of the world will approach]					
	A	<i>spizavyj</i>	„Zr-calo“,	<i>to jako gadaniem-</i>	<i>položij</i>	
	and	write.PTCP.PST.NOM.SG	“Zr-calo”	as guessing	assume.3SG	
	<i>7000 lět,</i>	<i>glagoljaj</i>	<i>sice:</i>	<i>Po domyslu</i>	<i>ubo</i>	<i>rěša_k</i>
	7000 years	speak.AP	so	by conjecture	for	decide.PST.3PL
	<i>mnozi_k,</i>	<i>glagoljušček:</i>		<i>poneže v- 6 dnej s-tvoren-</i>	<i>byst’ mir- sej,</i>	
	many.NOM	speak.PTCP.PRS.NOM.PL		because this world was created in 6 days		

<i>i skončavaemě ubo šestoj tysjašči byti i skončanie miru:</i>				
the end of the sixth millenium must also be the end of the world				
<i>sl-gaša_k</i>	<i>že</i>	<i>tii,</i>	<i>reče_j.</i>	
lie.PST.3PL	but	those	speak.AOR.3SG	
‘Guessing, the author of “Zr-calo” presumes 7000 years, saying: “Reasoning deductively, many had decided that, as the world had been created within six days, the end of the sixth millenium must also be the end of the world; yet they lied, he said [...]”’				
(Poslanie Gennadija Ioasafu)				

Over the time, the verbal forms that can be used as markers narrow down to the imperfective adverbial participle: In the 18th century, the verb *rešči* ‘to say’, ‘to speak’ falls out of use altogether; the adverbial participle *glagolja* ‘saying’ is used only in archaic texts to mark direct speech and eventually gets out of use completely. By that time, the adverbial participle *govorja* ‘speaking’ occurs not only with deictic adverbs, but can also be modified with a limited number of other adverbs. These adverbs show a semantic parallel to the adverbs used in French to form DSEs based on the infinitive *parler* ‘to speak’. At the same time, Russian DSEs based on the infinitive *skazat’* ‘to say’ are attested for the first time, which can be interpreted as a pattern replication from French (cf. Birzer 2013). The contact modes with French were both *literacy contact* and, to a more limited extent, *face-to-face contact*, so the bilingual persons very likely had receptive and productive skills in both oral and written form (the relevant literature often points out that French educators and correspondence in French were quite common in Russian upper-class circles, cf. Offord, Ryazanova-Clarke, Rjéoutski and Argent 2015 among others). Consequently, there is reason to assume that the bilingual persons were – at least to some extent – aware of the morphological and (morpho)syntactic di- and convergences in French and Russian, which possibly lead to a catalyst effect on the morphosyntactic level. In the 19th century, the type and token frequencies (graphs 2 and 3 respectively) for DSEs based on the adverbial participle increase, while the frequencies for infinitive-based DSEs stagnate (cf. Birzer 2012: 239-242 for a detailed discussion). We may deduce that the replication of the French pattern served as a catalyst for the expansion of the genuinely Slavic construction pattern.



Graph 2: Type frequencies for DSEs based on the adverbial participle or infinitive (data basis from the Russian National Corpus; cited after Birzer 2012: 242).



Graph 3: Token frequencies for DSEs based on the adverbial participle or infinitive (data basis from the Russian national corpus; cited after Birzer 2012: 241).

By analogy to the construction ADV + AP, the contextualising DSEs of the type ADJ.INST + LANGUAGE.INST + NON-FINITE VERBUM DICENDI and the quotative DSEs of the type LANGUAGE.INST + PROPER_NAME.GEN + NON-FINITE VERBUM DICENDI emerge in the 19th century.

4.1.2 Czech

Since Czech was not included in any of the previous studies on DSEs, this section is somewhat more comprehensive than those on the other languages.

Research on discourse markers and discourse connectives has quite a tradition in Czech linguistics (for a short outline of the research history cf. Rysová 2018: 41–44). With the rise of large annotated text corpora, corpus-based approaches especially to discourse connectives have moved into the focus of research (cf. Rysová 2012; Zikánová, Hajičová, Hladká, Jínová, Mírovský, Nedoluzhko, Poláková, Rysová, Rysová and Václ 2015; Poláková 2015; Rysová and Rysová 2015; Mírovský, Synková, Rysová and Poláková 2017).

Discourse connectives are divided into primary and secondary connectives (cf. Rysová and Rysová 2018). The latter include DSEs of the type ADV + *receno* (cf. Rysová 2012: 2804; Poláková 2015: 54).¹³ The distinction between primary and secondary connectives is made not solely on morphological grounds – petrified univerbations are considered to be primary connectives and “fixed phrases (both grammatically and lexically restricted) that are fully frozen [...] or that enable only a slight modification” (Rysová and Rysová 2014: 454; cf. also Danlos, Rysová, Rysová and Stede 2018: 73–74) to be secondary ones – but also takes into account the ability to “express the reference in the surface” (Rysová and Rysová 2014: 456; cf. also Rysová 2012: 2802), i.e. to syntactically conjunct two sentences or even clauses (instead of establishing reference to a preceding context of any syntactic size; for a list of all distinctive characteristics cf. Rysová and Rysová 2014: 458). Usage frequency is not explicitly mentioned as criterion, but it is striking that the core and periphery according to usage frequency coincides

¹³ Quite interestingly, Rysová’s (2014) study on the connective usage of verbs of saying does not consider the verb *říct* ‘say’.

with the status of primary vs. secondary discourse connective (cf. Rysová and Rysová 2018: 28).

There have been only a few investigations on the historical development of discourse markers and connectives: Based on data from historical dictionaries and the DIAKORP corpus, Doischer (2015) analyses the development of *vždyt* ‘after all’. In one of the introductory chapter Doischer discusses what distinguishes grammaticalization from pragmaticalization, but does not analyse the development of *vždyt* from this perspective. Rysová (2017) examines the historical development of some very frequent primary connectives in Czech and English (with some references to German) and discusses whether secondary connectives can develop into primary ones. The DSEs of the type ADV + *řečeno* are not explicitly addressed. Both questions are also examined from the perspective of grammaticalization theory; pragmaticalization as a further explanatory approach is not considered.

Let us now turn to our corpus-based study of the historical development of DSEs in Czech. The data stem from the following subcorpora of the Czech National Corpus: Diakorp6, KH Noviny and KH Dopisy as additional resources for the 19th century and Speeches.

Of the four languages discussed, Czech is the only one where the competition between the two non-finite verb forms exists already in the first attestations of both forms in discourse structuring function.¹⁴ In the 14th century, the present participle active (as predecessor of the adverbial participle, i.e. *transgressive* in Czech linguistic terminology) of *říct* ‘say’ occurs as marker of direct and indirect speech (25). In several cases the participle is modified by *tak(to)* ‘so’:

(25)	Old Czech				
	<i>O tom</i>	<i>také</i>	<i>píše</i>	<i>Nikodemus</i>	<i>v svém čtení</i>
	about this	also	write.PRS.3SG	Nikodemus.NOM	in his reader
	<i>a</i>	<i>tak</i>	<i>řka,</i>	<i>jež</i>	<i>dva syny Simeonova,</i>
	CONJ	so.ADV	say.PTCP.PRS.ACT.NOM.SG.M	COMP	two sons of Simon,
	<i>Karinus a Alerchius , jsta z mrtvých s Jezukristem vstala</i>				
	Karinus and Alerchius, resurrected from the dead together with Jesus Christ				
	‘Nikodemus also writes about this in his reader, saying so that two sons of Simon, Karinus and Alerchius, resurrected from the dead together with Jesus Christ ...’				
	(X: Pasionál muzejní (Muz III D 44) (R) (1350–1400))				

As in the other languages discussed, the modification by *tak(to)* is the point of departure towards the development of DSEs, as example (26) impressively shows: the adverb *zvlášče* ‘especially, particularly’ expresses the narrator’s conviction that the following quote is the actual wording and may thus be considered a predecessor of DSEs expressing speaker’s stance:

(26)	Old Czech				
	<i>Ježíš</i>	<i>matce</i>	<i>odpovídáše</i>	<i>a</i>	<i>zvlášče</i>
	Jesus.NOM	mother.DAT	answer.AOR.3SG	CONJ	in particular.ADV
	<i>takto</i>	<i>řka:</i>	<i>Matko má milá, ...</i>		
	so.ADV	say.PTCP.PRS.ACT.NOM.SG.M	my dear mother		
	‘Jesus gave an answer to his mother speaking in particular so: “My dear mother, ...” ’				
	(X: Pasionál muzejní (Muz III D 44) (R) (1350–1400))				

¹⁴ A favoring factor for the early attestations is the fact that the diachronic subcorpus of the Czech National Corpus is more comprehensive than the diachronic (sub)corpora that cover the period from the 14th–17th century for the other languages studied.

It is worth highlighting that representatives of other discourse structuring functions are also attested as early as in the 14th century. (27) is an example of reformulation, and (28–29) illustrate the first step towards the contemporary reformulation marker *česky řečeno* ‘in plain words (lit. speaking Czech)’. It is quite noteworthy that in example (26) the reformulation is integrated into the direct speech, although it provides the Czech equivalent to the original address and thus constitutes metatext. Quite often, the past participle passive still occurs with an auxiliary, i.e. as element of a composite final verb form (25–26).

(27)	Old Czech		
	<i>Ni v Čechách, ni u Moravě, jakžto učení mistři pravie,</i>		
	neither in Bohemia, nor in Moravia, as the learned masters say,		
	<i>ni v Rakúsiach, ni v Uhřiech, ni u Bavořiech ani v Rusiech,</i>		
	nor in Austria, nor in Hungary, nor in Bavaria, nor in Russia,		
	<i>ni u Polaniech, ni v Korutaniech – právěť vešde jeho jmě světie,</i>		
	nor in Poland, nor in Carinthia – everywhere his holy name is to reign		
	<i>krátčě</i>	<i>řkúce,</i>	<i>po všem světě –</i>
	more_ shortly.ADV	say.PTCP.PRS.SG.F	on the whole world –
	<i>nikdiež jemu nenie rovně ...</i>		
	nowhere is an equal one to him		
	‘Neither in Bohemia, nor in Moravia, as the learned masters say, nor in Austria, nor in Hungary, nor in Bavaria, nor in Russia, nor in Poland, nor in Carinthia - his holy name is to reign everywhere, in short, all over the world - there is none equal to him.’		
	(X: Mastičkář muzejní (R) (1350-1400))		

(28)	Old Czech			
	<i>Tehda křesťené [...] jemu jmě vzděli Přispořený.</i>			
	then the Christians gave him the name Přispořený [endowed with wealth]			
	<i>To</i>	<i>je</i>	<i>českým</i>	<i>obyčjem</i>
	DET.SG.N	AUX.PRS.3SG	Czech.ADJ.INST	custom.INST
	<i>řečeno</i>		<i>Sbožný.</i>	
	say.PTCP.PST.PASS.SG.N		Sbožný	
	‘Then the Christians [...] gave him the name Přispořený. In Czech words, that is Sbožný.’			
	(X: Pasionál muzejní (Muz III D 44) (R) (1350-1400))			

(29)	Old Czech				
	<i>[...] oni řechu:</i>	<i>“Rabi,</i>	<i>to</i>	<i>jest</i>	<i>řečeno,</i>
	they said	Rabbi.VOC	DET.SG.N	AUX.PRS.3SG	say.PTCP.PST.PASS.SG.N
	<i>mistře,</i>	<i>kde bydlíš?”</i>			
	master.VOC	where do you live?			
	‘[...] they said: “Rabbi, that is master, where do you live?” ’				
	(X: Život Krista Pána (UK XVII A 9) (R))				

Instances of textual reference are attested as well (30).

(30)	Old Czech			
	<i>A to proto,</i>	<i>jakž</i>	<i>jest</i>	<i>dřeve</i>
	and this is for the reason	how	AUX.PRS.3SG	earlier.ADV
	<i>řečeno,</i>			
	<i>aby za ny všichni světi před bohem prosili.</i>			
	say.PTCP.PST.PASS.SG.N	that for them all saints would beg before God		

	‘And this is so for the reason, as has been said before, that for them all saints would beg before God.’
	(X: Pasionál muzejní (Muz III D 44) (R) (1350-1400))

From the 15th century onwards the present participle active of *řéci / říct* ‘say’ is attested only as marker of speech rendering, in some cases modified by the adverb *tak(to)* ‘so’. Its usage frequency declines sharply in the 17th c. Instead, the present participle active of *říkat*, the imperfective aspectual partner of *řéci / říct* ‘say’, occurs in a discourse-structuring function for the first time in the 16th century. It marks speaker’s stance (31); in the 17th century it also functions as a quotation marker of the type *s.PREP PROPER_NAME.INST říkajíc* ‘speaking in the words of person X (lit. speaking with person X)’ (32):¹⁵

(31)	Middle Czech		
	[...] <i>toliko pochovávati mrtvé jest</i>	<i>jako</i>	<i>říkajíc</i>
	to bury so many dead is	how	say.PTCP.PRS.ACT.F/N
	<i>práce nesnesitelná, [...]</i>		
	an unbearable work		
	‘Burying so many dead is, so to say, an unbearable task.’		
	(Hostounský, Baltazar: Obrácení pohanův v Jáponě (1585))		

(32)	Middle Czech		
	<i>Těšte se, ó, ženy počestné, jistou přítomnosti boží,</i>		<i>říkajíce</i>
	Rejoice, oh, honorable women, in the sure presence of God,		say.PTCP.PRS.ACT.PL
	<i>s</i>	<i>Davidem</i>	<i>z</i>
	with	David.INST	from
			<i>žalmu 73:</i>
			psalm.GEN 73
	<i>Podlé rady své ved’ mne, Bože, a potom v slávu přijmeš mne.</i>		
	According to your counsel lead me, oh God, and then you will receive me in glory.		
	‘Rejoice, oh, honourable women, in the sure presence of God, saying with David in Psalm 73: According to thy counsel lead me, oh God, and then you will receive me in glory.’		
	(Žalanský-Phaeton, Havel: O ctných manželkách těhotných a rodičkách křesťanských (Strah. knih., BT VIII 6) (1615))		

The distribution of functions in the 16th and 17th centuries results from the development of verbal aspect into a grammatical category:¹⁶ the rendered speech acts are part of the narration; they are episodic (and thus also semelfactive) and are therefore marked by the participle of the perfective verb *řéci / říct* ‘say’. In the 18th century, the present participle of *řéci / říct* ‘say’ as marker of rendered speech is out of use. This correlates with the development of binding rules for punctuation. Speaker’s stance (hedging, i.e. the usage of a downtoning formulation in order to avoid loss of face on the speaker’s or addressee’s side, forms an aspect of speaker’s stance in a broader sense) and references to the sources of quotations, on the other hand, do not form part of the narration, i.e. they are not episodic, and are therefore conveyed with imperfective *říkajíc*. From the 18th century on, the present participle respectively transgressive of *říkat* is used exclusively as hedge in collocation with *tak* ‘so’, but up to and even in the 19th century it

¹⁵ *říkajíc* is used as a proxy for all inflectional forms of the present participle, since the hedging function crystallizes on *říkajíc* later on.

¹⁶ According to Kosek (2014: 158), the development of aspect towards a grammatical category has its climax in the 15th c.

featured different inflectional forms. The crystallization to the petrified (and sometimes unverbated) *tak říkajíc* ‘so to say (lit. saying so)’ is a result of the 20th century.¹⁷

The functions attested for *řečeno* ‘said’ in the 14–16th centuries remain stable over the centuries. In the 19th century, as in the other languages studied, the number of DSEs of the type *adv + řečeno* increases.

If we now compare the Czech structures with those in its contact language German, the following parallels stand out: in German and in Czech, the DSE with hedging function is formed with the equivalent adverbs Czech *tak* ‘so’ and GER *so* ‘so’, respectively. The non-finite verb form used – the transgressive in Czech and the infinitive in German – differs from the inflectional form used in both languages for the (semi)productive formation pattern: both languages use the participle passive for the latter purpose. The German hedge *sozusagen* ‘so to speak’ results from the unverbation of the three constituents *so*, *zu* and *sagen*,¹⁸ in Czech, the unverbation *takříkajíc* ‘so to speak’ exists alongside the separate spelling. These parallels in form and function make it very likely that the Czech DSEs are pattern replications of the German models.

4.1.3 Croatian

In modern Croatian, the imperfective adverbial participle *govoreći* ‘speaking’ occurs in two major contexts: a) as element of the DSE types ADV + AP and ADJ.INST + LANGUAGE.INST + NON-FINITE VERBUM DICENDI and b) in unmodified form as marker of direct speech (22; for this and the following cf. Birzer 2018a).

(33) Croatian

*Pokraj nje bila su prevrnutá dječja kolica, a djevojčica.NOM je.AUX.3SG
plakala.LPTCP.SG govoreći.AP: “Što si napravio mojoj mami!?”*

¹⁷In the syn2020 corpus only *tak říkajíc* is attested, the other two inflectional forms of the transgressive *tak + říkaje / říkajice* do not occur.

¹⁸The “Deutsches Textarchiv” (<https://www.deutschestextarchiv.de/>) contains texts of the New High German period from the 17th to the 20th century. Our query showed that *sozusagen* ‘so to say’ – still in the separate spelling – was used in hedging (1) and reformulation (2) function already in the 17th century. The first instances of unverbation are documented in the 1840s, cf. (3), well before the first evidence in Czech and still at a time of intensive (literacy) contact between Czech and German. The last evidence for the separate spelling in German dates back to 1910.

(1) *Diethelm fühlte, was es heißt, sich in schwankender oder gar in verzweifelter Lage zu befinden, da muß man sich so zu sagen übers Ohr hauen lassen und thun als ob nichts geschehen wäre, nur um Aufsehen und genauere Nachforschung zu vermeiden.*

‘Diethelm felt what it means to be in a shaky or even desperate situation, where you have to let yourself be taken for a ride, so to speak, and act as if nothing had happened, just to avoid a fuss and a more detailed investigation.’

(Auerbach, Berthold: *Die Geschichte des Diethelm von Buchenberg*. In *Deutscher Novellenschatz*. Hrsg. von Paul Heyse und Hermann Kurz. Bd. 7. 2. Aufl. Berlin. [1910]).

Die Zitationen sind so übernommen, wie sie im Korpus gegeben wurden. Sie können gerne der Einheitlichkeit halber angepasst oder gelöscht werden.

(2) *In beiden Ständen der Mensch nur gedacht als Person, so weit er besitzt, eine Darstellung gewissermaßen des flüchtigen Lebensgeistes, der aus der Scholle aufsteigt, die Repräsentation, so zu sagen, des Ackerduftes.*

‘In both states man is thought of only as a person, given he has possessions, a representation, in a sense, of the ephemeral spirit of life that rises from the soil, the representation, so to speak, of the scent of the field.’

(Immermann, Karl: *Der Carneval und die Sommambüle*. In *Deutscher Novellenschatz*. Hrsg. von Paul Heyse und Hermann Kurz. Bd. 5. 2. Aufl. Berlin. [1910], 139–273).

(3) *Solche psychische Dispositionen kommen unzweifelhaft angeboren, und namentlich angeerbt, sozusagen häufig eben als Träger der Heredität des Irreseins, vor [...]*

‘Such mental dispositions are undoubtedly innate, and especially inherited, and so to speak occur frequently as carriers of the heredity of insanity.’ (Griesinger, Wilhelm: *Die Pathologie und Therapie der psychischen Krankheiten, für Ärzte und Studierende*. Stuttgart. 1845).

‘Next to her the pram was turned over, and the girl cried, saying: “What have you done to my mama!?”’

(HNK_v25 subcorpus, gs20050702ck42973)

Until the middle of the 17th century, prototypical *govoreći*.AP ‘speaking’ and its (still existing participial) variants is used exclusively as a marker of direct and indirect speech as well as quotations. It occurs a) in the construction with two *verba dicendi* (34) and b) as a syntactically independent parenthesis within or at the end of the rendered speech or quotation (35), i.e. the status of the secondary predicate gets lost gradually:

(34) Middle Croatian

... kamenovahu Stiepana zovećega i govorećega:
lapidate.IMPERF.3PL Stiepan.ACC cry.PTCP.ACT.ACC and speak.PTCP.ACT.ACC
Gñe Jesuse, primi duha moga.

Lord Jesus, accept my soul.

‘... they lapidated Stiepan, who was crying and speaking: “Lord Jesus, accept my soul.”’

(Kašić, Dj 7:58/Acts 7:59)

(35) Middle Croatian

On ju tišeći tisućkrat i tisuć milo pritiskajući
he her console.AP thousand_times and thousand heartily press.AP
celunu: - Nemoj se - govoreći - dušice moja, rascviljevati.
kiss.3SG do not speak.AP my darling, moan

‘Consoling and a thousand and thousand times pressing her heartily, he kissed her: “Don’t,” saying, “my darling, moan.”’

(HNK_klasici subcorpus, Zoranić P. Planine. 1569).

In the 18th century, *govoreći* retains the described functions, but it also sporadically occurs in the construction type ADV + AP (36). In addition, the constructions with the past participle passive *rečeno* ‘said’, still as a predicate nominal with auxiliary, are attested (37). They express speaker’s stance:

(36) Middle Croatian

... skoro sasvim ništa ne valjade. Al po duši naški
... nearly in general nothing turns around. but by soul.DAT ADJ.POSS.DAT
govoreći, umrit ćemo, ...
speak.AP we will die

‘... nearly in general nothing turns around. But speaking according to our soul, we will die ...’

(HNK_klasici subcorpus, Relković, M. A. Satir iliti divji čovik. 1762)

(37) Middle Croatian

Pođe s ovoga svita kralj slovinski isto vrime udari na Ljutovida neizbrojena vojska cesara Ludovika, kojoj ne mogući Ljutovid odoliti, pobiže u srbsku zemlju i bi primljen od vojvode srbskoga u dvor svoj; ali se Ljutovid ukaza nepoznan i žestoko neharan, jer pogubivši svoga dobročasnika učini se gospodar od njegova grada i svega bogatstva. Pravo.ADV je.AUX.3SG rečeno.PTCP.PASS.SG: “Ne čini dobra nepoznanu, da te zlo ne nađe”.

‘The Slavic king [i.e. Borna – S.B.] had departed from this world and at that time uncountable armed forces of Emperor Ludovik made an attack against Ljutovid, and Ljutovid, unable to resist them, fled to the Serbian lands and was received by the Serbian

duke in his court; but Ljutovid turned out unthankful and terribly disrespectful, because having killed his benefactor he made himself sovereign of his city and all riches. Rightly is said: “Don’t do good to an unknown person, so evil will not find you.”

(HNK_klasici subcorpus, Kačić-Miošić, A. Razgovor ugodni naroda slovinskoga. 1759)

The structural parallel with German DSEs of the type ADV + PTCP.PST.PASS like *ehrlich / kurz gesagt* ‘honestly / briefly said’ etc. as well as the increasing influence of German on Croatian in the 18th century suggest a pattern replication from German (cf., among others, Rammelmeyer 1975; Štebih 2003; Golubović 2007; Jernej, Glovacki-Bernardi and Sujoldžić 2012; on the role model of Czech for replications from German in other Slavonic languages, among them Croatian, cf. Turk and Sesar 2003). In the 19th century, *rečeno* lost its predicative status and became the preferred non-finite verb form for forming DSEs. By analogy with the DSEs of the type ADV + NON-FINITE VERBUM DICENDI, the contextualising DSEs of the type ADJ.INST + LANGUAGE.INST + NON-FINITE VERBUM DICENDI and the citation-marking DSEs of the type LANGUAGE.INST + PROPER_NOUN.GEN + NON-FINITE VERBUM DICENDI evolved. Unmodified *govoreći.AP* ‘speaking’ becomes restricted to marking direct speech. In parallel with Czech (and in contrast to Russian and Polish), the replicated construction is the (semi)productive pattern for forming DSEs in Croatian (cf. Chart 1). This is likely to be the starting point for a crystallization of functions: Constructions with *rečeno* function as DSEs, while *govoreći* prototypically occurs as a marker of direct speech.

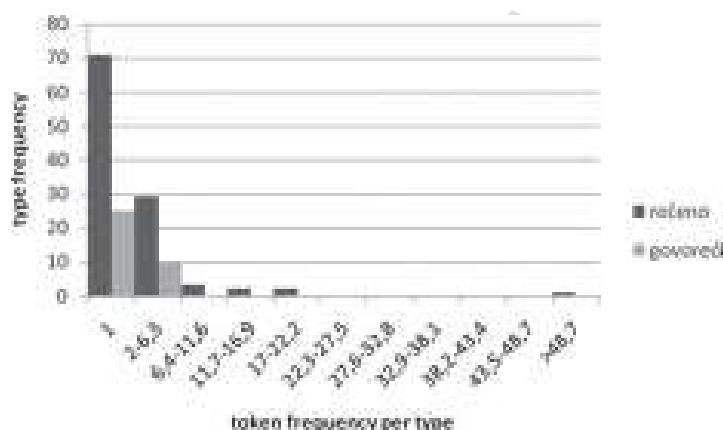


Chart 1. Histogram of type and token frequencies in the HNK25 corpus for the DSEs based on *govoreći* and *rečeno* (cited from Birzer 2018a: 229).¹⁹

Of all the language contact situations investigated in our study, the contact between Croatian and German probably has the second-highest degree of intensity (after Czech, cf. Thomas 2003), since it must be assumed that, at least in the cities, relatively broad sections of the population had receptive and productive skills in oral and written German (cf. Jernej et al. Jernej, Glovacki-Bernardi and Sujoldžić 2012: 333–336). The Croatian pattern replication also testifies to the awareness of morphosyntactic con- and divergences. Even a catalyst effect in the crystallization of the functions cannot be excluded: Croatian speakers may have been aware that the German construction ADV + *gesagt*.PTCP.PASS ‘said’ does not function as marker of rendered speech, which possibly drove the specialization of *govoreći.AP* ‘speaking’ in a marker of direct speech.

4.1.4 Polish

¹⁹ The classes were created automatically on the basis of the data of the more infrequent *govoreći*.

The historical development of Polish DSEs (cf. Birzer 2017b) is similar to that in Russian. Constructions for marking rendered speech and quotations are also the first to be attested in Polish. Congruent structures can be observed foremost in religious Latin-Polish parallel texts. In earlier periods of Polish, it is often impossible to distinguish between direct and reported speech,²⁰ while Latin uses the AcI as specialized construction to mark reported speech.

In the period of macaronism, Polish replicates the syntactic pattern of the AcI – the AcI may be formed with both Latin and / or Polish lexical material – so *oratio recta* and *obliqua* can now be distinguished (37; for this and the following, cf. Birzer 2015b):

- (37) Middle Polish
- | | | | | | |
|-------------------|---------------------------------|--------------------|----------------|------------|-----------------|
| ... <i>przeto</i> | <i>Orpheus</i> | <i>powiedział,</i> | [<i>Jovem</i> | <i>być</i> | <i>masculum</i> |
| therefore | Orpheus.NOM | say.PST.3SG.M | Jupiter.NOM | be.INF | man.ACC |
| <i>et</i> | <i>feminam</i>] _{AcI} | | | | |
| and | woman.ACC | | | | |
- ‘... therefore Orpheus said that Jupiter was at the same time man and woman.’
(Górnicki. 1566. Dworzanin Polski.)

Latin insertions also occur in the quotative function, which develops from speech marking. If the original language of the quotation is Latin, this is not surprising, but Latin insertions also occur in contexts in which the information about the quoted author is so sparse that the original language of the quotation cannot be determined. Therefore, it must be assumed that Latin insertions serve to distinguish between the actual narration (in Polish) and quotation (or indirect speech – which is also a kind of quotation – in Latin) and can thus be described as functional code switches (cf. Riehl 2004: 23; Pahta and Nurmi 2011). The AcI got out of use at the turn of the 18th to the 19th centuries, i.e. at a time when punctuation was becoming increasingly standardised and direct speech as well as verbatim quotations could be clearly identified by means of punctuation (cf. Birzer 2015b).

Latin lexical material was also used to form DSEs of the type ADV + AP. Gradually, the lexical material became integrated morphologically and finally, some individual DSEs were replicated using the Polish equivalents for the borrowed Latin material (e.g. *mówiąc in particulari* > *partykularnie mówiąc* > *ściśle(j) / konkretnie(j) / dokładnie(j) mówiąc* ‘(more) exactly speaking’). Replacing the borrowed material by its Polish equivalents is a case of matter replication. The replication acts as a catalyst for the consolidation of the genuinely Slavic construction ADV + AP.

Thus, in the development of Polish discourse structures the replication of both matter (in the DSEs) and pattern (the AcI in quotative function) took place (cf. Birzer 2017b).

The contact between Polish and Latin has to be classified as intensive literacy contact. Two products of language contact testify to the awareness of morphosyntactic con- and divergences typical of intensive literacy contact: firstly, the replication of the AcI and secondly the morphological integration of the borrowed Latin matter into the Polish orthographic and morphological system, which reflects the Latin morpheme boundaries.²¹

4.2 Cross-linguistic analysis: the role of language contact and extralinguistic factors

²⁰ In earlier periods of Polish, complementizers may mark both direct and reported speech. If the respective speech act contains a pronoun and the context makes clear that the pronoun shift typical for reported speech has taken place, it is clear that we are dealing with an instance of reported speech. However, in many instances the context does not allow to establish whether the (third person) pronoun in the speech act is the result of pronoun shift or not.

²¹ This is supported by the colloquial matter replication *baks* ‘dollar’ in Modern Russian, which results from a language contact situation of considerably less intensity and does not reflect the morphological segmentation of *buck-s*.

Comparing the historical development of DSEs in Czech, Croatian, Polish and Russian, a number of similarities and differences can be found. All four languages share the following characteristics: a) the contemporary DSEs developed from the same construction for speech marking, which occurs frequently in religious texts. b) They all feature a construction type based on the imperfective adverbial participle, which competes with another construction type based on another non-finite verb form. c) Language contact exerted influence on the historical development of the DSEs based on non-finite *verba dicendi*: It produced a catalytic effect that led to the firm rooting of one of the competing construction types into the linguistic system. On the smaller scale, however, the effects of language contact differ greatly.

In Polish, language contact leads to the borrowing of lexical material which fills the adverb position of the genuinely Slavic construction ADV + AP.IPF (and to the temporarily used replication of the AcI), whereas in Russian as well as Czech and Croatian it results in the borrowing of a discourse-structuring construction from French (based on the infinitive) and German (based on the past participle passive) respectively. This is remarkable since Matras (2007b: 68) concludes from his typological data that pattern replication occurs with connective constructions, to which many DSEs with pragmatic functions belong, whereas matter replication occurs only with discourse-structuring particles. The intensity of language contact explains why the borrowed construction forms only isolated lexicalised representatives in Russian, but is (semi-)productive in Czech and Croatian: In Russia, only the educated classes had knowledge of French and *literacy contact* was probably more pronounced than *face-to-face contact*, whereas in Bohemia and Moravia as well as in Croatia, as historical-political entities of the Habsburg Empire, contact with German took place face-to-face in many areas of everyday life, but also in the form of *literacy contact* especially in the official domain (cf. Jernej et al. – Jernej, Glovacki-Bernardi and Sujoldžić 2012; Newerkla 2022: 159–162), in Bohemia and Moravia even to a greater extent than in Croatia (cf. Nuckols 2003: 117; Thomas 2003: 206–207). Additionally, Czech and Croatian witnessed literacy contact via mutual translations of mainly literary works (cf. Stehlík 2019: 57–59) and, to a smaller extent, face-to-face contact due to Czech migration mainly to Slavonia (cf. Stehlík 2019). This possibly served as additional catalyst for the consolidation of DSEs based on the past participle passive in both languages.

Specifically for Croatian, the establishment of the replicated construction with the past participle passive as the (semi-)productive one possibly served as the starting point for the preservation of unmodified *govoreći* as marker of direct speech, which consequently resulted in the crystallization of functions (marking of direct speech vs. discourse structure).

At the same time, our data show that the replication of patterns cannot be explained by a gap in the linguistic system of the replicating language, since a genuinely Slavic construction for marking discourse-structuring functions is present in all four languages.

Rather, extralinguistic factors, namely stylistic and literary fashions, play a role in the consolidation of the constructions studied. In the 19th century, the number of DSEs grew rapidly in all four languages. At the same time, the new narrative technique of so-called “experienced speech” (*erlebte Rede*) emerges, which “jene Stellen in einem schriftlichen oder mündlichen Text [bezeichnet], die in einer gegebenen Rede die Frage aufkommen lassen, wer da “eigentlich” spricht (denkt/wahrnimmt). Also: E.R. [erlebte Rede – S.B.] als Irritation der Redeinstanz und in deren Folge: E.R. als Form der Interferenz von Primär- und Sekundärtext [refers to those passages in a written or oral text that raise the question of who is “actually” speaking (thinking/perceiving). Thus: experienced speech as an irritation of the speech instance and in its wake: experienced speech as a form of interference of primary and secondary text – translation S.B.]” (Hodel 2001: 49). Several (secondary) characteristics of experienced speech have a direct relation to one of the functions performed by the DSEs under study:

- (a) The “wertungsmässige [evaluative – translation S.B.]” orientation (cf. Hodel 2001: 45) implies the differing evaluation of one and the same situation and parallels the discourse-

structuring function of speaker's stance, prototypically expressed by the construction ADV + NON-FINITE VERBUM DICENDI.

- (b) Contextualization within a discourse and its prototypical construction ADJ.INST + LANGUAGE.INST + NON-FINITE VERBUM DICENDI can be regarded as an “antidote” to the irritation caused by the unmarked interference of primary and secondary text. The contextualizing construction indicates the primary text parts of which have been transferred into the secondary text.
- (c) The spectrum of the so-called “zitierende (citatnaja) Rede reicht von einzelnen Lexemen bis zu ‘subjektiven Redemassiven’ [quotative (citatnaja) speech ranges from single lexemes to ‘subjective speech massives’ – translation S.B.]” (Hodel 2001: 32). In quotative speech the “Enunziator EN [ist] eine von der Sprecherinstanz SI zu unterscheidende Äußerungsinstanz, die anhand unterschiedener sprachlicher Manifestationen bestimmbar ist [enunciator EN [is] the instance of utterance that has to be distinguished from the speaker instance SI; the enunciator can be identified by their characteristic linguistic instantiations – translation S.B.]” (Hodel 2001: 39). The same distinction is made by the construction LANGUAGE.INST + proper noun.GEN + NON-FINITE VERBUM DICENDI, which serves as marker of quotations (and their sources).

The described parallels between the narrative technique of experienced speech and the DSEs studied make it highly likely that the emergence of experienced speech caused the above-mentioned rise in type and token frequency of the DSEs.

The borrowing of Latin lexical material for the Polish DSEs can be explained in a similar way: at the time of their replication, “macaronism” was in vogue, so again an extralinguistic factor lead to the consolidation of a linguistic structure.

So far, only sociopragmatic changes (cf. for example the model in Kuße 2009: 50) have been considered as extralinguistic factors, but, to the best of our knowledge, literary devices such as narrative techniques have not been addressed as extralinguistic factors fostering language change. Interestingly, sociopragmatic change has oral communication as point of departure, yet narrative techniques used in the *belles lettres* presuppose written media. Therefore, one may hypothesize that face-to-face and literacy contact respectively are influenced by different extralinguistic factors.

5 Grammaticalization, lexicalization and pragmaticalization

In recent years some studies appeared that treated Slavic DSEs of different morphological form from the perspective of pragmaticalization (cf. Pekelis 2020 on the Russian particle *že*, Benigni 2020 on Russian *uže* ‘already’, Bobrik 2021 on the Russian farewell formula *ščastlivo* and Vinogradova 2023 on constructions containing the preposition *po* ‘by, through’; for Polish Pastuch 2022 discusses the discourse function of the imperative *śluchaj* ‘hear’ and Kubicka 2017, 2018 even treats DSEs based on *mówiąc*.AP ‘speaking’, however focussing on issues of synonymy in a very general sense). However, none of the papers contributes to the further development of the theory, i.e., specifically to the question of how the development of DSEs relates to the theoretical concepts of grammaticalization, lexicalization, and pragmaticalization.

5.1 State of research

Considering grammaticalization, lexicalization and pragmaticalization in order to assess the historical development of DSEs is motivated as follows: two of the three constructions that form DSEs, namely the constructions based on the adverbial participle or infinitive respectively, behave morphosyntactically like the prepositions that developed from adverbial participles (e.g. Russian *spustja* ‘later’ or Croatian *zahvaljujući* ‘thanks to’): their co-reference with the first

argument of the matrix clause is lost. Since the development of the aforementioned prepositions is a case of grammaticalization (Gxn) (cf. Birzer 2010: 243–250), it is obvious to verify whether the development of the DSEs also constitutes a Gxn process. Another fact makes the verification advisable as well:

For all non-finite verb forms that serve as DSE bases, by default the speaker becomes the covert (and, with the past participle passive, also the demoted) first argument of the *verbum dicendi*. This can be interpreted as a case of subjectification (cf. Traugott and Dasher 2002), whose relation to Gxn is described by Traugott as follows:

subjectification and intersubjectification are independent of grammaticalization. However, since grammaticalization involves the development of markers of speaker attitude towards the ideational component and towards textual connectivity (among many other things), there is inevitably a close interaction between grammaticalization and subjectification. (Traugott 2010: 60).

This raises several theoretical issues:

- a) Is the development of the described DSEs a case of Gxn, since the speaker as the covert first argument of the non-finite verbal bases indicates subjectification, and subjectification in turn closely interacts with Gxn?
- b) Does the subjectification occur independently of a Gxn process and is the development of the DSEs consequently to be understood as a lexicalization process (Lxn process)? An argument in favour is Birzer's (2010: 261–264) finding that the semantic processes in Gxn and Lxn partially correspond, as the development of adverbs and prepositions out of adverbial participles shows.
- c) Finally, we need to pay attention to Diewald's (2011) position that pragmaticalization (Pxn; for a short history of pragmaticalization research cf. Heine 2013: 1217–1220) is a specific subtype of Gxn, since the discussion about the link between subjectification and Gxn has gained new momentum through this approach. Drawing on Diewald (2011), Beijering (2012) identifies overlaps between Gxn, Pxn and Lxn, which leads her to propose the following model:²²

A unified account of language change centred on the notion of a composite change. A composite change is composed of formal reanalysis and semantic reinterpretation, accompanying primitive changes at different linguistic levels (phonology, morphology, syntax, semantics, discourse / pragmatics), and the concomitants of these changes. These (micro-)changes cluster with respect to their converging properties (=interfaces), but stand out in case they are diverging properties (=unique properties). [. . .] It has been shown that primitive changes tend to form clusters in terms of reduction or expansion (e.g. semantic bleaching versus enrichment, morphological fusion versus separation) or no change at all. These clusters may correspond to changes traditionally labelled 'grammaticalization', 'lexicalization' or 'pragmaticalization', but changes may also cluster in alternative ways. (Beijering 2012: 225).

- d) The development of DSEs is an instance of co-optation, as proposed by Heine (2013). The cooptation hypothesis posits that DSEs arise through a process of cooptation followed by a grammaticalization process.

Cooptation is defined as follows:

Cooptation is a fully productive operation whereby a chunk of sentence grammar, such as a word, a phrase, a reduced clause, a full clause, or some other piece of text, is deployed for use on the metatextual level of discourse processing, thereby turning into a thetical. Its functions are

²² For a detailed discussion of the approaches of Diewald and Beijering, see Birzer (2012a, 2015a).

determined by the situation of discourse, serving (a) to overcome constraints imposed by linearization in structuring texts, (b) to provide the source of information, (c) to place a text in a wider perspective, e.g. by elaborating, proffering an explanation, a comment or supplementary information, (d) to describe the attitudes of the speaker, and/or (e) to interact with the hearer. (Heine and Kaltenböck 2021: 7)

Discussing the cooptation approach is important because three of the six English comment clause types analyzed by Heine and Kaltenböck (2021) as instances of cooptation are structurally similar to the Slavic DSEs under investigation in this paper.

Heine and Kaltenböck's type (v) in English incorporates an “-ing clause (*speaking as a layman*)” (Heine and Kaltenböck 2021: 2). The -ing form is highly polyfunctional and serves as an adverbial participle (cf. Kortmann 1995), making type (v) the structural equivalent of Slavic DSEs based on the adverbial participle.

Type (vi) “take[s] an -ed clause (*stated bluntly*)” (Heine and Kaltenböck 2021: 2). In this context, -ed has the function of participle passive, and type (vi) may thus be regarded as the structural equivalent of the Czech and Croatian DSEs based on the passive participle in -no.

Type (iv) “take[s] a to-infinitive (*to be honest*)” (Heine and Kaltenböck 2021: 2) and displays the lowest degree of structural convergence. Of the four languages analyzed, only Russian features DSEs based on the infinitive *skazat'* ‘say’.

Heine identifies spontaneity, scope, syntax, semantics, morphophonology, prosody, and word order as features distinguishing cooptation from grammaticalization (Heine 2013: 1222–1223). Since our findings for syntax, semantics, morphophonology, and prosody coincide with Heine's observations, we will take a closer look only at spontaneity, scope, and word order.

We will begin with scope and word order. Heine claims that “grammaticalization tends to lead to a restriction in the (semantic-pragmatic) scope of the unit undergoing the process, whereas cooptation entails an increase in scope” (Heine 2013: 1222). He further suggests that “[u]nits coopted as theticals [...] tend to be freed of constraints on placement [...], even if there usually is some preferred position” (Heine 2013: 1223). However, Birzer (2015) shows for the Polish and Russian DSEs under investigation that they seem to be free of placement constraints only on first sight; in fact, each function of these polyfunctional items requires a specific placement and correlates with a specific scope that may range from one wordform in syntactic contact position up to utterances/paragraphs or even whole texts. Thus, these two features apply to the DSEs under consideration only in a limited manner.

Heine states that “[c]ooptation, by contrast [to grammaticalization – S.B.], is spontaneous rather than gradual; it can take place any time and in any situation” (Kaltenböck, Heine and Kuteva ~~et al.~~ 2011: 874–875). The formulation seems to imply that apart from completely spontaneous instances of cooptation, there may also occur “more gradual” ones. Unfortunately, it remains unclear which “degree of gradualness” is still acceptable for cooptation and which degree is indicative of grammaticalization, which is always gradual. This issue is crucial for the DSEs under investigation, as our data from four languages shows that the development of the DSEs is gradual and follows a certain path. Only the increase in type frequency in the 19th century is leap-like. We may describe this leap also as a rise in productivity. Heine describes extension as a process where “[t]he thetical is used frequently and in more contexts” (Heine 2013: 1223). This rather implies a more frequent usage of one – lexicalized – DSE in various contexts and not a rise in construction productivity. Instead, the increased productivity of our DSEs may be considered an instance of extension, which Heine attributes to the grammaticalization following the cooptation. As our data shows, the DSEs of the type ADV + PTCP.PASS enter Czech and Croatian rather “spontaneously” and are most likely the result of

language contact with German. Thus, the theoretical question arises whether cooptation may be contact-induced and in which ways it differs from contact-induced grammaticalization.²³

5.2 Application to the DSEs under investigation

Lehmann's (2002) parameters of grammaticalization serve as the basis for the following discussion. In case the DSEs with pragmatic, contextualising or quotative functions differ in their behaviour with respect to a given parameter, they will be discussed individually with regard to that parameter. In all other cases they are treated together.

Since Diewald's (2011) parameters of pragmaticalization are variations of Lehmann's parameters, they will be dealt with – if relevant to our question – following Lehmann's corresponding parameter. The insights gained will then be related to the properties of Lxn processes, with Brinton and Traugott (2005) in particular serving as a basis, and to Heine's cooptation hypothesis.

5.2.1 Integrity

Lehmann (2002: 112) distinguishes between semantic and phonological integrity; in the course of a Gxn process, semantic erosion and possibly phonological attrition takes place. With regard to semantic erosion, one may state that with all DSEs based on *verba dicendi* the semantic component 'oral articulation' erodes. At the same time, however, the new semantic component 'the speaker' emerges, caused by the specification of the speaker as the covert first argument of the non-finite *verbum dicendi*.

This is in line with Heine's finding that "coopted units [...] have non-restrictive or metacommunicative meaning that is not part of the sentence meaning" (Heine 2012: 1222), but one may question whether successive desemanticization takes place, i.e. the loss of "most or all of the lexical-conceptual meaning it [the DSE – S.B.] may have had in favor of discourse-organizing functions" (Heine 2013: 1223). The latter is typical of the grammaticilization process following cooptation, but is more pronounced with the hedging or filler function of some DSEs and less pronounced with discourse structuring functions.

Croatian unmodified *govoreći* deserves explicit mention, as it retains the semantic component 'oral articulation'. Although co-reference with the first argument of the matrix sentence is temporarily no more obligatory even at a rather early stage of development (38), the speaker does not by default become the covered first argument of *govoreći*, as can be seen from the contemporary example (39):

- (38) Croatian
Govoreći tako: da će svih zgubiti ki ne htiše, kako on reče.
 speak.AP so COMPL that he will kill how he say.AOR.3SG
 'Saying so: that he will kill [another possible interpretation would be *izgubiti*
 'lose'] all who do not want, as he said.'
 (HNK_Marulić, Judita)

- (39)
- | | | | |
|--|------------------|---------------------------------------|-----------|
| Croatian | | | |
| <i>Pokraj nje bila su prevrnutu dječja kolica,</i> | <i>a</i> | <i>djevojčica</i> | <i>je</i> |
| next to her the pram was turned over | and | girl.NOM | AUX.3SG |
| <i>plakala</i> | <i>govoreći:</i> | <i>“Što si napravio mojoj mami!?”</i> | |
| cry.PTCP.SG.F | speak.AP | What have you done to my mama!? | |

²³ Castle's (2021) dissertation on structural transfer in Czech-English language contact bears the verb *co-option* in the title, but in the text body *co-option* is neither defined nor used. Judging from the content, co-optioning is equaled to structural replication in the sense of Heine and Kuteva (2008).

‘Next to her the pram was turned over, and the girl cried, saying: “What have you done to my mama!?”’

(HNK_v25, gs20050702ck42973)

If quotative *govoreći* were to be considered a case of Gxn, the restoration of co-reference and thus the return to the starting point of development would contradict the proposed unidirectionality of Gxn processes.

Phonological reduction can be observed only for DSEs with pragmatic functions and, as a rule, it occurs only with a few highly frequent DSEs. Reduction takes place in two steps. In the first step, the verbal element, i.e. the adverbial participle (40) or passive participle (41), is omitted:

(40) Polish

... *każdy człowiek ma elementy – geny, hormony, cechy psychiczne – męskie i żeńskie lub, ogólniej, aktywne i pasywne ...*

‘... every person has elements – genes, hormones, mental traits – male and female or, more generally, active and passive ones ...’

(NKJP, J. Prokopiuk. 2007. *Dusza ludzka – oś świata.*)

(41) Croatian

Prof. dr. Vinko Čandrlić sa suradnicima predlaže da se Jadranska autocesta vodi od graničnog prijelaza Rupe do Grobničkog polja gdje bi se uključila u Autocestu Rijeka-Zagreb, a od sadašnjeg čvora Vrata (Fužine) da se napravi spoj do Žute Lokve i dalje trasom autoceste Zagreb-Split kroz Liku kao što se i službeno predlaže. Tako bismo imali gotovu trasiranu autocestu što vodi od Trsta do Atene. Time bi se i Gorski kotar približio moru. Ukratko, Jadransko-jonska autocesta „zaobišla“ bi grad Rijeku i ne bi devastirala primorska turistička područja.

‘Prof. Dr. Vinko Čandrlić and his colleagues propose that the Adriatic Motorway be led from the border town of Rupe to Grobničko polje, where it would join the Rijeka-Zagreb Motorway, and from the current hub Vrata (Fužine) that a connection be made to Žuta Lokva and then along the Zagreb-Split Motorway through Lika, as is also proposed by the services. So we would have a ready-made motorway route from Trieste to Athens. This way the Gorski kotar region would also become closer to the sea. **In short**, the Adriatic-Ionian motorway would bypass the city of Rijeka and would not devastate the coastal tourist areas.’

(HNK_v30, text CW080199908271601hr).

With still a smaller number of DSEs an additional reduction of the non-verbal element is attested, e.g. Russian *voobščē* > *vooščē* ‘generally’ (the reduced form is even noted down in the oral sub-corpus of the Russian National Corpus).

If we were dealing with a Gxn process, phonological erosion should occur with all (or a large part of) DSEs. However, this expectation is not fulfilled: Firstly, the omission of the verbal element exclusively takes place in DSEs based on the adverbial participle or the participle passive, but elision does not occur with DSEs employing the infinitive. Secondly, the elision of the verbal element co-occurs only with some specific functions of certain DSEs, but is not their universal trait.

Rather, the selective elision of the adverbial participle or participle passive may be considered an instance of clipping in the sense of Brinton and Traugott (2005: 40–41), which is typical for lexicalization.

5.2.2 Paradigmatization

In all our object languages, a construction for marking speech acts, typically containing an (adverbial) participle, forms the point of departure for the development of DSEs based on non-finite *verba dicendi*. This construction may be regarded as genuinely Slavic. In its original function as speech act marker, various *verba dicendi* could fill the slot of the adverbial participle, but in the course of development only one verb lemma remained in each object language to fill the slot, namely Russian *govorja* ‘speaking’, Croatian *govoreći* ‘speaking’ and Czech *říkajíc* ‘speaking’. One can argue whether the restriction to one verb lemma constitutes a narrowing of the paradigm in the sense of Gxn theory. If it were not for the alleged spontaneity of co-optation, there were more reason to regard the process as the co-optation of a specific inflectional form of a specific verb lexeme. The situation in Polish is slightly different: in addition to the semi-productive construction with the imperfective adverbial participle *mówiąc* ‘speaking’, a few collocations based on the construction with the perfective adverbial participle *powiedziałwszy* ‘having said’ have been preserved. The latter all have pragmatic functions. In Polish, the perfective adverbial participle as such is generally in decline. Thus, the petrification of these collocations can be seen as instances of lexicalization. A parallel can be discerned to Russian adverbs that constitute petrified adverbial participial forms based on unproductive suffixes (cf. Birzer 2010: 114–186). Polish even features a limited number of doublets such as *prawdę mówiąc*.AP.IPF / *powiedziałwszy*.AP.PFV ‘frankly speaking (lit. speaking / having said the truth)’, which speaks against grammaticalization at work.

Another argument against grammaticalization is the fact that in each object language, language contact resulted in the competition between two functionally equivalent construction types or, in the case of Polish, in constructional doublets (*generalnie* vs. *ogólnie mówiąc* ‘generally speaking’). This can be understood as a relatively recent formal widening of the paradigm.

It should also be noted that – in contrast to the pragmatic function – the contextualising and the quotative function are both bound to one specific morphosyntactic construction, namely LANGUAGE.INSTR + PROPER.NAME.GEN + AP. Overall, paradigmaticity is more pronounced in DSEs with contextualising or quotative functions than in DSEs with pragmatic functions.

5.2.3 Paradigmatic variability

Lehmann (²2002: 124) links the parameter of paradigmatic variability with the measurement criterion of obligatoriness, which Diewald (2011: 368) in turn splits up into intra-linguistic and communicative obligatoriness. She considers the latter to be typical of Pxn processes and

intended to capture the fact that many categories are obligatory in the sense that they have to be realised in the relevant position, i.e., the speaker cannot leave them unspecified if s/he does not want to produce incorrect utterances, but the choice among the paradigmatic members of the category is not determined by language internal features but by the communicative intentions of the speaker (Diewald 2011: 368).

Intralinguistic obligatoriness, which usually finds its expression at the level of morphology or morphosyntax, does not apply to any single group of DSEs, since the insertion of a DSE is neither morphologically nor morphosyntactically necessary.

Communicative paradigmaticity and thus obligatoriness requires a more differentiated discussion, as explained in Birzer (2012a): although it is possible to form grammatically correct sentences without using a DSE, the individual groups of DSEs nevertheless exhibit varying degrees of communicative obligatoriness: two subgroups of DSEs with pragmatic functions, namely DSEs with a connective or topic management function, make information structures explicit; thereby they promote information processing and thus contribute to communicative

success. A third subgroup of DSEs with pragmatic function, namely the DSEs with hedging function, contribute to communicative success, as hedging may help to prevent a loss of face. DSEs with contextualising or quotative information also contribute to the explicitness of information structures as they mark that the speaker relates part of their utterance to another discourse or has inserted an element from another source into their utterance; however, this does not necessarily enhance information processing. Finally, DSEs with the pragmatic function of filler words cannot be ascribed any of the properties mentioned.

As Birzer (2015a) illustrates with Russian *voobščę govorja* 'generally speaking' and Polish *ogólnie mówiąc* 'generally speaking', many DSEs incorporate a whole range of pragmatic functions which differ in communicative obligatoriness and thus with regard to communicative paradigmaticity. Since Diewald (2011) developed her parameters on analysing modal particles, i.e. a group of homogeneous and monofunctional elements, she did not encounter the problem of polyfunctional elements and accordingly does not address it.

If the functions of one DSE vary with regard to obligatoriness, it is difficult to attribute the given DSE a general degree of G_{xn} or P_{xn}.

Varying obligatoriness does not contradict co-optation. Consequently, one might raise the question whether the DSEs under investigation have undergone co-optation, but grammaticalization is only at an initial stage. Additionally, one might discuss whether co-optation is necessarily followed by grammaticalization (Heine 2013 and Heine and Kaltenböck 2021 do not discuss the option of co-optation without grammaticalization).

5.2.4 Scope, bondedness and syntagmatic variability

Since these three parameters interact, we will treat them together.

Regarding bondedness, it suffices to mention that in general DSEs are neither bound nor do they have clitic status.

Concerning scope and syntagmatic variability, DSEs with pragmatic functions display a behaviour pattern different from those with contextualising or quotative functions.

At first sight, DSEs with multiple pragmatic functions seem to be unrestricted in their syntagmatic variability. However, as the case study (Birzer 2015a) shows, this is not true; rather, polyfunctional DSEs feature a well-defined syntagmatic position and scope for each individual function. The filler word function is an exception, since filler words can occur in any position, yet they do not have any scope at all. In fact, it is the missing scope that makes absolute syntagmatic variability possible in the first place.

DSEs with contextualising or quotative function are usually placed in preposition to the element they scope over. Their extension of scope, however, is somewhat diffuse, as it cannot extend beyond a sentence but may vary below this threshold.

To summarize, the syntagmatic variability of the filler word function corresponds with Heine's word order feature characteristic for co-optation, whereas the syntagmatic behaviour of all other functions is rather typical of grammaticalization.

6. Conclusion

In conclusion, one may state that the DSEs with contextualising or quotative functions predominantly show a behaviour that is indicative of a G_{xn} process, whereas the DSEs with pragmatic functions display a behaviour typical of L_{xn} with regard to integrity and paradigmaticity. Their paradigmatic and syntagmatic variability and the leap-like increase of DSEs in the 18th and 19th centuries neatly fits in with co-optation, but rather not with grammaticalization (Heine 2013 claims the latter to necessarily follow co-optation).

This leads us to reconsider an assumption already made in Birzer (2015a), namely that the development of the DSEs based on non-finite *verba dicendi* as entire group cannot be encompassed by the analytical apparatus of Gxn, Lxn and Pxn research available so far. Instead it should be interpreted with the help of Beijering's model of composite change (2012). This model allows to form one grammaticalization cluster comprising the DSEs with contextualising or quotative functions. We may then assume that the grammaticalization cluster has an interface with the co-optation cluster formed by DSEs with pragmatic functions, whose individual functions comply with a varying range of Gxn parameters. Finally, the co-optation cluster also provides an interface to the lexicalization cluster characterized predominantly by paradigmatic variability and phonetic erosion.

References

- Beijering, Karin. 2012. *Expressions of epistemic modality in Mainland Scandinavian. A study into the lexicalization-grammaticalization-pragmaticalization interface*. (Groningen Dissertations in Linguistics, 106). Groningen. <http://dissertations.ub.rug.nl/faculties/arts/2012/k.beijering/?pLanguage=en>
- Benigni, Valentina. 2020. Èto uže meloči. From phasal adverb to intensifier: The pragmaticalization of uže. *Studi Slavistici*, XVII. 131–155. https://doi.org/10.13128/STUDI_SLAVIS-9115.
- Berger, Tilman. 2009. Einige Bemerkungen zum tschechischen Absentiv. In Tilman Berger et al. (eds.) *Von grammatischen Kategorien und sprachlichen Weltbildern – Die Slavia von der Sprachgeschichte bis zur Politsprache. Festschrift für Daniel Weiss zum 60. Geburtstag*, 9–28. München/Wien: Otto Sagner.
- Berger, Tilman. 2014. The convergence of Czech and German between the years 900 and 1500. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in contact-induced language change. Language families, typological resemblance, and perceived similarity*. 184–198. Berlin/Boston: DeGruyter.
- Bilková, Jana, Denisa Knobová & Sára Podojilová. 2021. K Emotivnosti Diskurzivních Markerů ve Spontánní Mluvené Komunikaci. *Stylistyka* 30. 113–24. <https://doi.org/10.25167/Stylistyka30.2021.5>.
- Birzer, Sandra (accepted). Formal unity and functional diversity: A corpus-linguistic approach to Russian and Polish adverbial participles with the meaning 'speaking' between discourse and grammar. In Mirjam Fried & Eva Lehečková (eds.), *Conjunctions vs. contextualizers*. John Benjamins. (working title).
- Birzer, Sandra. 2010. *Russkoe deepričastie: processy grammatikalizacii i leksikalizacii*. [The Russian adverbial participle: processes of grammaticalization and lexicalization.] Munich: Otto Sagner.
- Birzer, Sandra. 2012a. From subject to subjectivity: Russian discourse structuring elements based on the adverbial participle *govorja* 'speaking'. *Russian Linguistics* 36(3). 221–249.
- Birzer, Sandra. 2012b. Von der (in)direkten Rede zum Kommentar. *Verba dicendi* in griechischen und altkirchenslavischen Konstruktionen der Redewiedergabe. In Björn Hansen (ed.), *Diachrone Aspekte slavischer Sprachen*, 197–209. München/Berlin: Otto Sagner.
- Birzer, Sandra. 2013. Constructions based on the Russian adverbial participle *govorja* and their function in discourse and grammar. In Sebastian Kempen, Monika Wingender, Norbert Franz & Miranda Jakiša (eds.), *Deutsche Beiträge zum 15. Internationalen Slavistenkongress Minsk 2013*, 67–77. München: Otto Sagner.
- Birzer, Sandra. 2015a. Generally speaking, connectivity and conversation management combined. The functions of Russian *voobščee govorja* and Polish *ogólnie mówiąc* 'generally speaking'. *Russian Linguistics* 39(1). 81–115.
- Birzer, Sandra. 2015b. Kosmetik auf den Lippen der Szlachta. *Verba dicendi* und der AcI in der Geschichte des Polnischen. Vortrag an der Universität Freiburg, 15.05.2015.
- Birzer, Sandra. 2017a. Assessing the role of pattern and matter replication in the development of Polish discourse structuring elements based on non-finite *verba dicendi*. Part 1 – Synchronic state. *Zeitschrift für Slavische Philologie* 73(1). 159–185.

- Birzer, Sandra. 2017b. Assessing the role of pattern and matter replication in the development of Polish discourse structuring elements based on non-finite *verba dicendi*. Part 2 – Historical development. *Zeitschrift für Slavische Philologie* 73(2). 397–440.
- Birzer, Sandra. 2018a. Historical development and contemporary usage of discourse structuring elements based on *verba dicendi* in Croatian. In Jasmina Grković-Major, Björn Hansen & Barbara Sonnenhauser (eds.), *Diachronic Slavonic syntax. The interplay between internal development, language contact and metalinguistic factors*, 209–258. Berlin/Boston: De Gruyter Mouton.
- Birzer, Sandra. 2018b. Nimm zwei? Konkurrierende Konstruktionen und Sprachkontakt in der Slavia am Beispiel diskursstrukturierender Elemente auf der Basis infinitiver *verba dicendi*. *Die Welt der Slawen* 63(1), 1–39.
- Birzer, Sandra. 2018c. Semantic explications of Polish DSEs with pragmatic functions and assignment of their non-verbal elements to semantic hyperclasses [Data set]. Otto-Friedrich-Universität Bamberg. <https://doi.org/10.48564/unibafid-3gvr9-kzr75>
- Bobrik, Marina. 2021. Pragmaticalization mechanisms in the history of the Russian farewell formula *sčastlivo! Voprosy jazykoznanija* 2021(1). 70–83.
- Boguslavskij, Igor' M. 1996. *Sfera dejstvija leksičeskix edinic*. [The scope of lexical units]. Moskva: Škola "Jazyki Russkoj Kul'tury".
- Boguslavskij, Igor' M. & Leonid L. Iomdin. 2009. Moscow Semantic School. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung. Band 1*, 813–827. Berlin/New York: DeGruyter.
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge UP.
- Brinton, Laurel J. 2017. *The evolution of pragmatic markers in English: Pathways of change*. Cambridge: Cambridge University Press.
- Castle, Chloe Michelle. 2021. The co-option of grammatical resources between languages: A focus on English and Czech. <https://digital.library.adelaide.edu.au/dspace/handle/2440/134191> (last access date October 1, 2023).
- Danlos, Laurence, Kateřina Rysová, Magdaléna Rysová & Manfred Stede. 2018. Primary and secondary discourse connectives: Definitions and lexicons. *Dialogue & Discourse* 9(1). 50–78.
- Degand, Liesbeth & Jacqueline Evers-Vermeul. Year? Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. *Journal of Historical Pragmatics* 16(1). 59–85.
- Dehé, Nicole & Yordana Kavalova (eds.). 2007. *Parentheticals*. Amsterdam: John Benjamins.
- del Saz, Rubio M. & Bruce Fraser. 2003. Reformulation in English. <http://people.bu.edu/bfraser/> (last access October 1, 2023).
- Dickey, Stephen M. 2013. See, now they vanish: Third-person perfect auxiliaries in Old and Middle Czech. *Journal of Slavic Linguistics* 21(1). 77–121.
- Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2). 365–390.
- Eufe, Rembert. 2013. *Verba Dicendi im Gebrauch. Dicere in der Vita Corbiniani und ihrer Überarbeitung*. In Sabine de Knop et al. (ed.). *Konstruktionsgrammatik in den romanischen Sprachen*. 261–297. Frankfurt am Main u. a.: Lang.
- Evans, Robert J. W. 2004. Language and state building: The case of the Habsburg monarchy. *Austrian History Yearbook* 35. 1–24.
- Fischer, Kerstin. 2006. Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In Kerstin Fischer (ed.), *Approaches to discourse particles*. 1–20. Amsterdam: Elsevier.
- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6(2). 167–190.
- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31(7). 931–952.
- Fraser, Bruce. 2006. Towards a theory of discourse markers. In Kerstin Fischer (ed.), *Approaches to discourse particles*. 189–204. Amsterdam: Elsevier.
- Fried, Mirjam & Jan-Ola Östman. 2004. Construction Grammar: a thumbnail sketch. In Mirjam Fried & Jan-Ola Östman (eds.), *Construction Grammar in a cross-language perspective*. 11–86. Amsterdam: John Benjamins.

- Fried, Mirjam. 2010. A Construction Grammar approach to grammatical change. In Björn Hansen & Jasmina Grković-Major (eds.), *Diachronic Slavonic syntax. Gradual changes in focus*. 53–62. München/Berlin/Wien: Otto Sagner.
- Grochowski, Maciej, Anna Kisiel & Magdalena Żabowska. 2014. *Słownik gniazdowy partykuł polskich*. [Wordnest dictionary of Polish particles]. Kraków: Polska Akademia Umiejętności.
- Grochowski, Maciej. 1983. Metatekstowa interpretacja parentezy. [The metatextual interpretation of parenthesis.] In Teresa Dobrzyńska & Elżbieta Janus (eds.), *Tekst i zdanie. Tom studiów*. [Text and sentence. A volume of studies.], 247–258. Wrocław: Ossoliński.
- Grochowski, Maciej. 1986a. O metapredykatywnej funkcji niektórych wyrażen partykułowo-przysłówkowych w strukturze tekstu. [On the metapredicative function of some particle-adverbial expressions in the text structure.] In Teresa Dobrzyńska (ed.), *Teoria tekstu. Zbiór studiów*. [Theory of text. A collection of studies.], 139–148. Wrocław et al., Ossoliński.
- Grochowski, Maciej. 1986b. *Polskie partykuły. Składnia, semantyka, leksykografia*. [Polish particles. Syntax, semantics, lexicography.]. Wrocław.
- Grochowski, Maciej. 1998. *Funktionswörter im Polnischen*. Oldenburg.
- Hansen, Björn & Sandra Birzer. 2012. The Yiddish modal system between Germanic and Slavonic. A case study on limits of contact induced grammaticalization. In Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical replication and borrowability in language contact*, 427–464. Berlin: Mouton de Gruyter.
- Hansen, Björn. 2010. Mood in Polish. In Björn Rothstein & Rolf Thieroff (eds.), *Mood in the languages of Europe*, 342–357. Amsterdam: John Benjamins.
- Heine, Bernd & Gunther Kaltenböck. 2021. From clause to discourse marker: On the development of comment clauses. *Language Sciences* 87. 1–17.
- Heine, Bernd & Tania Kuteva. 2008. Constraints on contact-induced linguistic change. *Journal of Language Contact* 2(1). 1–34.
- Heine, Bernd & Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3). 529–572.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: CUP.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55(4). 813–855.
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6). 1205–1247.
- Hladký, Ladislav (ed.). 2019. *Czech relations with the nations and countries of Southeastern Europe*. Zagreb: Srednja Europa.
- Hodel, Robert. 2001. *Erlebte Rede in der russischen Literatur. Vom Sentimentalismus zum Sozialistischen Realismus*. Band 1. Frankfurt am Main: Peter Lang.
- Imo, Wolfgang & Susanne Günthner (eds.). 2006. *Konstruktionen in der Interaktion*. Berlin: de Gruyter.
- Imo, Wolfgang. 2007. *Construction Grammar und Gesprochene-Sprache-Forschung: Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen: Niemeyer.
- Itani, Reiko. 1995. A relevance-based analysis of Lakoffian hedges: *sort of, a typical and technically*. In *UCL Working Papers in Linguistics*, 7. 87–105. <http://www.phon.ucl.ac.uk/publications/WPL/95papers/ITANI.pdf> (last access October 1, 2023).
- Jernej, Mirna, Zrinjka Glovacki-Bernardi & Anita Sujoldžić. 2012. Multilingualism in northwestern part of Croatia during Habsburg rule. *Jezikoslovlje* 13(2). 327–50.
- Kaltenböck, Gunther, Bernd Heine & Tania Kuteva. 2011. On thetical grammar. *Studies in Language* 35(4). 852–897.
- Kolyaseva, Alena. 2022. The divergent paths of pragmaticalization: The case of the Russian particles *tipa* and *vrode*. *Journal of Pragmatics* 201. 181–96.
- Kortmann, Bernd. 1995. Adverbial participial clauses in English. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective*. 189–238. Berlin: De Gruyter.
- Kosek, Pavel. 2014. *Historická mluvnice češtiny. Vol. 1*. Brno: Masarykova univerzita.
- Kubicka, Emilia. 2017. Jak Mówimy, Jakoś Mówiąc? Formalne i Semantyczne Właściwości Adverbialnych Uzupełnień Quasi-Imiesłowowego *Mówiąc*. [How do we speak, speaking somehow? Formal and semantic properties of the adverbial complements of the quasi-adverbial participle *mówiąc* 'speaking']. *Saying LingVaria* 23. 99–113.

- Kuße, Holger. 2009. Kleinsprachenlinguistik. Zur Typologie von Kleinsprachen und den Aufgaben ihrer linguistischen Beschreibung am Beispiel slavischer Kleinsprachen. In Christian Prunisch (ed.), *Konzeptualisierung und Status kleiner Kulturen: Beiträge zur gleichnamigen Konferenz in Dresden vom 3. bis 6. März 2008*, 41–60. München: Otto Sagner.
- Kyseláková, Jana. 2013. *Aorist ve staročeské bibli*. [The aorist in the Old Czech Bible]. Brno: BA thesis. 2013. https://is.muni.cz/th/kjmx0/Bakalarska_prace_Kyselakova.pdf (last access October 1, 2023).
- Lauwers, Peter, Gudrun Vanderbauwhede, & Stijn Verleyen (eds.). 2012. *Pragmatic markers and grammaticalization: Lessons from false friends*. Amsterdam/Philadelphia: John Benjamins.
- Lehmann, Christian. 2002. *Thoughts on grammaticalization*. <http://www.christianlehmann.eu/publ/ASSidUE09.pdf> ((last access October 1, 2023).
- Magdaléna Rysová & Kateřina Rysová. 2014. The centre and periphery of discourse connectives. *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing*, 452–59. Phuket, Thailand: Department of Linguistics, Chulalongkorn University.
- Marác, László. 2012. Multilingualism in the transleithanian part of the Austro-Hungarian empire (1867-1918): Policy and practice. *Jezikoslovlje* 13. 269–298.
- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2). 281–331.
- Matras, Yaron. 2007. Contact, connectivity and language evolution. In Jochen Rehbein, Christiane Hohenstein & Lukas Pietsch (eds.), *Connectivity in grammar and discourse*, 51–74. Amsterdam: John Benjamins.
- Matras, Yaron. 2007. The borrowability of grammatical categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–74. Berlin: Mouton de Gruyter.
- Mendoza, Imke. 2009. Anaphorische Mittel: Konnexion. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung. Band 1*, 982–990. Berlin/New York: DeGruyter.
- Mírovský, Jiří, Pavlína Synková, Magdaléna Rysová & Lucie Poláková. 2017. CzeDLex – A lexicon of Czech discourse connectives. *The Prague Bulletin of Mathematical Linguistics* 109(1). 61–91.
- Mosegaard-Hansen, Maj-Britt. 2006. A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French *toujours*). In Kerstin Fischer (ed.), *Approaches to discourse particles*, 21–41. Amsterdam: Elsevier.
- Myers-Scotton, Carol. 2002. *Contact linguistics. Bilingual encounters and grammatical outcomes*. Oxford: Oxford UP.
- Newerkla, Stefan Michael. 2022. Reconstructing multilingualism in the Habsburg state: Lessons learnt and implications for historical sociolinguistics. *Sociolinguistica* 36(1–2). 151–67.
- Nikolaeva, Tat'jana M. 2008. *Neparadigmatičeskaja lingvistika: istorija „bluždajuščix častic“*. [Non-paradigmatic linguistics: the history of “stray particles”]. Moskva: Jazyki slavjanskix kul'tur.
- Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14(2). 177–20.
- Nuckols, Mark E. 2003. An evaluation of German-Croatian contact. Working paper. Ohio State University. Department of Linguistics. 100–119. <https://kb.osu.edu/handle/1811/81023> (last access October 1, 2023).
- Nyan, T. 2006. From procedural meaning to processing requirement. In Kerstin Fischer (ed.), *Approaches to discourse particles*. Amsterdam. 167–188.
- Offord, Derek, Lara Ryazanova-Clarke, Vladislav Rjéoutski & Gesine Argent (eds.). 2015. *French and Russian in Imperial Russia. Vol. 1. Language use among the Russian elite*. Edinburgh: Edinburgh UP.
- Pahta, Päivi & Arja Nurmi. 2012. Multilingual discourse in the domain of religion in medieval and early modern England: A corpus approach to research on historical code-switching. In Herbert Schendel & Laura Wright (eds.), *Code-switching in early English*. 219–252. Berlin/Boston: DeGruyter.
- Pastuch, Magdalena. 2022. O rozwoju dyskursywnej funkcji form rozkazujących (Na przykładzie *śluchaj*). [On the development of a discursive function of imperative forms (by the example *śluchaj*)]. *LingVaria* 17(2). 143–62.

- Pekelis, Olga E. 2020. Ob odnom slučae pragmatikalizaciji v ruskom jazyke: mikrodiachroničeskoe issledovanie časticy že v sostave voprosa. [About one case of pragmaticalization: a microdiachronic study of the particle *že* in a question clause.] *Slověne* 9(1). 340–61.
- Rabus, Achim. 2013. *Die Rolle des Sprachkontakts für die slavischen (Standard-)Sprachen (unter besonderer Berücksichtigung des innerslavischen Kontakts)*. Freiburg i. Br. (Ms.) <https://app.box.com/s/9ck4rh5d7fyjci4yyyx8> (last access October 1, 2023).
- Radanović-Kocić, Vesna. 1996. The placement of Serbo-Croatian clitics: A prosodic approach. In Aaron Halpern & Arnold Zwicky (eds.), *Approaching second: second position clitics and related phenomena*, 429–445. Stanford: CSLI Publications.
- Rathmayr, Renate. 1985. *Die russischen Partikeln als Pragmalexeme*. München: Otto Sagner.
- Reinkowski, Ljilijana. 2001. *Syntaktischer Wandel im Kroatischen am Beispiel der Enklitika*. München: Otto Sagner.
- Reinkowski, Ljilijana. 2008. Zum Erlernen der Syntax im Kroatischen: Die Problematik der Enklitika. In Biljana Golubović & Jochen Raecke (Hrsg.), *Bosnisch, Kroatisch, Serbisch als Fremdsprachen an den Universitäten der Welt*. (Die Welt der Slawen. Sammelbände 31). 131–141. München: Otto Sagner.
- Riehl, Claudia. 2004. *Sprachkontaktforschung. Eine Einführung*. Tübingen: Niemeyer.
- Riemer, Nich. 2006. Reductive paraphrase and meaning: a critique of Wierzbickian semantics. *Linguistics and Philosophy* 29. 347–379.
- Rysová, Magaléna. 2017. Discourse connectives: From historical origin to present-day development. In Katrin Menzel, Ekaterina Lapshinova-Koltunski & Kerstin Kunz (eds.), *New perspectives on cohesion and coherence: Implications for translation*, 11–34. Berlin: Language Science Press.
- Rysová, Magdaléna & Kateřina Rysová. 2015. Secondary connectives in the Prague dependency treebank. In Joakim Nivre & Eva Hajičová (eds.), *Proceedings of the third International conference on dependency linguistics (Depling 2015)*, 291–99. Uppsala: Uppsala University.
- Rysová, Magdaléna & Kateřina Rysová. 2018. Primary and secondary discourse connectives: Constraints and preferences. *Journal of Pragmatics* 130. 16–32.
- Rysová, Magdaléna. 2012. Alternative lexicalizations of discourse connectives in Czech. In *Proceedings of the eight International conference on language resources and evaluation (LREC'12)*, 2800–2807. Istanbul.
- Rysová, Magdaléna. 2014. Verbs of saying with a textual connecting function in the Prague discourse treebank. In *Proceedings of the ninth International conference on language resources and evaluation (LREC'14)*. 930–35. Reykjavik.
- Rysová, Magdaléna. 2018. *Diskurzivní konektory v češtině: od centra k periférii*. [Discourse connectors in Czech: from the center to the periphery.] Praha: Ústav formální a aplikované lingvistiky.
- Raxilina, Ekaterina V. (red.). 2010. *Grammatika konstrukcij*. [Construction Grammar]. Moskva: Azbukovnik.
- Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin: Mouton de Gruyter.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge UP.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collocations: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Stehlik, Petr. 2019. Czech-Croatian relations. In Ladislav Hladký (ed.), *Czech relations with the nations and countries of Southeastern Europe*, 49–81. Zagreb: Srednja Europa.
- Thomas, George. 2003. Puristic attitudes to German phraseological and syntactic calques in the Slavic languages of the former Habsburg empire. *Canadian Slavonic Papers* 45(1–2). 201–25.
- Thomason, Sarah & Terence Kaufmann. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Traugott, Elizabeth Closs & Richard Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge UP.
- Traugott, Elizabeth Closs. 2010. Revisiting subjectification and intersubjectification. In Kristin Davidse, Lieven Vandelandotte & Herbert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 29–70. Berlin: De Gruyter Mouton.
- Uryson, Elena V. 2008. Sintaksičeskie svojstva sojuzov i teorija valentnostej. [The syntactic properties of conjunctions and valency theory.] In Galina Kustova, Raisa Rozina & Aleksandr Bondarenko (eds.), *Dinamičeskie modeli. Slovo. Predloženie. Tekst. Sbornik statej v čest' E. V. Padučevoj*.

- [*Dynamic patterns. Word. Sentence. Text. A Festschrift for E. V. Padučeva.*]. 809–824. Moskva: Jazyki slavjanskoj kul'tury.
- Vinogradova, Ekaterina N. 2023. Grammaticalization, lexicalization, and pragmaticalization: Russian constructions with the preposition PO. *Voprosy jazykoznanija* 2023(1). 54–87.
- Volin, Jan, Lenka Weingartová & Oliver Niebuhr. 2017. The prosody of the Czech discourse marker 'jasně': An analysis of forms and functions. *Phonetica* 73(3–4). 314–37.
- Weiss, Daniel. 2011. MTT meets Construction Grammar: The treatment of argument structure. In Igor Boguslavsky & Leo Wanner (eds.), *Proceedings of the 5th International conference on Meaning–Text Theory. Barcelona, september 8–9, 2011*, 295–305. Barcelona.
- Wiemer, Björn & Björn Hansen. 2012. Assessing the range of contact-induced grammaticalization in Slavonic. In Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical replication and borrowability in language contact*, 67–155. Berlin: Mouton de Gruyter.
- Wierzbicka, Anna. 1971. Metatekst w tekście. [Metatext in the text]. In Maria Renata Mayenowa (ed.), *O spójności tekstu (Z dziejów form artystycznych w literaturze polskiej) [On text coherence. (From the history of artistic forms in Polish literature)]*. 21, 105–121. Wrocław: Ossolineum.
- Zikánová, Šárka, Eva Hajičová, Barbora Vidová-Hladká, Pavlína Jínová, Jiří Mírovský, Anna Nėdoluzko, Lucie Poláková, Kateřina Rysová, Magdalena Rysová, & Jan Václ. 2015. *Discourse and coherence: From the sentence structure to relations in text*. Prague: Ústav formální a aplikované lingvistiky.
- Zumstein, Mirjam. 2012. Syntaktische Aktionalitätsmarkierungen im altkirchenslavischen analytischen Passiv – Bemerkungen zur Wortfolge. In Björn Hansen (ed.), *Diachrone Aspekte slavischer Sprachen*, 259–269. München/Berlin: Otto Sagner.

Abbreviations

CF – clitic form of pronouns

FF – full form of pronouns