

# Capturing an oxymoron in the wild: Directive subordination in Slavic

Imke Mendoza,<sup>a</sup> Barbara Sonnenhauser,<sup>b</sup> Björn Wiemer<sup>c</sup>

<sup>a</sup> Universität Salzburg, Salzburg, Austria <imke.mendoza@plus.ac.at>

<sup>b</sup> Ludwig-Maximilians-Universität, Munich, Germany <barbara.sonnenhauser@lmu.de>

<sup>c</sup> Johannes-Gutenberg-Universität, Mainz, Germany <wiemberb@uni-mainz.de>

Loss of illocution, the presence of a clause-initial connective and word order are usually taken to be indicators of structural asymmetries in clause combining. In this exploratory study, we aim to operationalize the relation between word order and clause type, using Russian, Polish, and Slovene as representatives of the three major Slavic branches. On the assumption that clause-initial function words may indicate subordination, we analyze the distribution of the presumably unequivocal complementizers *że* (Pol.), *čto* (Ru.) and *da* (Slv.), and compare them with Pol. *niech*, Ru. *pust'* and Slv. *naj*. These elements function as illocutionary force indicating devices for directive speech acts, but at the same time show complementizer-like properties when introducing a clause that follows a clause containing a complement-taking predicate. Using corpus data from two different diachronic stages we try to establish diachronic and cross-linguistic patterns that provide information on possible links between word order and the 'complementizerhood' of a clause-initial element. Our findings reveal that the concepts of word order and connective, i.e. the very concepts, that are often used for diagnosing subordination, seem to be ill-defined and need to be reconsidered on the basis of thorough empirical research.

**KEYWORDS:** subordination, illocutionary force, directives, complementizers, word order, Slovene, Polish, Russian.

## 1. Diagnosing subordination

One of the central assumptions concerning clause-combining and, more specifically, complementation, is the existence of an opposition between main and subordinate clauses. However, it has often been conceded that the difference may not be categorial, but rather a matter of degree (e.g. Lehmann 1988, Weiss 1989, Raible 1992, Verstraete 2007, Gast & Diessel 2012, Diessel 2015), or that it is organized by way of 'canonical' and 'non-canonical' subordination (Axel-Tober 2012). Once we accept gradience, a variety of criteria, or diagnostics, are called for that are supposedly capable of disclosing structural asymmetries in clause-combining. However, very often such criteria are taken for granted, or they are not sufficiently considered in their interaction. This

carries the danger that notorious analytical problems, in particular when dealing with naturalistic data, remain unnoticed, or are glossed over.

Among indicators of subordination, we find presumably universal properties like shifts of person-deictic expressions and loss of illocutionary force (Verstraete 2005 and references therein). Both criteria are difficult to test, at least on corpus data, either because often no good diagnostics can be applied (e.g. for the loss of illocutionary force), or because the relevant diagnostic contexts are not easy to come by (e.g. speech reports in which shifts between first, second and third person may be identified).<sup>1</sup> Other criteria rely on language-specific properties, such as tense shifts (*consecutio temporum*), word order, or the presence of a clause-initial connective in a non-first conjunct within a pair or a sequence of clauses. The latter two properties also reveal inherent problems. Thus, although different word order patterns have often been regarded as robust indicators of subordination (cf. Kortmann 1997: 61, 364 and contributions on Celtic, Germanic and Greek in Siewierska 1998), these patterns have the potential of signaling subordination only in languages that tend to employ word order for the coding of grammatical (i.e. syntactic) relations. Moreover, the concept of word order itself is not as straightforward as it seems (see §§2.3, (4-5)). Reliance on clause-initial connectives, in turn, is based on the assumption that these unequivocally mark subordination. This assumption is problematic since clause-initial connectives are not necessarily indicative of subordination, but they may fulfill the same function as in syntactically and illocutionary independent clauses, and, conversely, subordination does not necessarily require an overt indication; instead, the link may be simply asyndetic (see §§2.1-2.2). A case in point is the marking of directive illocution: if the illocutionary function of some marker agrees with (some element in) the immediately preceding context, how can we show whether the illocutionary force of the clause has been lost after the integration in a neighboring clause? Structural asymmetry needs then to be confirmed by other criteria.

Slavic languages provide an ideal playground to demonstrate all these challenges. They lack rules of *consecutio temporum* as well as distinct ordering patterns concerning the major constituents S, V, O in main and subordinate clauses (as for the latter cf. Siewierska & Uhlířová 1998). This leaves us with clause-initial connectives as possible indicators of embedding. However, many clauses containing such elements systematically defy a clear categorization: clause-initial function words often behave in a way that qualifies them as possible complementizers, while the clauses which they introduce show properties that run counter to long-standing assumptions about subordination. This is particularly

true for clauses introduced by units that originate from directive-optative expressions, such as Polish *niech*, Russian *pust'*, Slovene *naj* (see §2.1).

We thus put to the test word order and the occurrence of clause-initial connectives as those properties associated with subordination that seem easiest to check due to their manifestation in linear ordering. Simultaneously, we expand the concept of word order beyond the sequence of the major constituents S, O, V. We focus on Polish, Russian and Slovene as representatives of the three major branches of Slavic (West, East, and South). As a testing ground we use the aforementioned directive-optative expressions (Pol. *niech*, Russ. *pust'*, Slv. *naj*) occurring in clause-initial position, since these cases of 'directive subordination' are a particular challenge to syntactic analysis. We examine the distributional properties of these units on a strictly empirical basis, also against the background of units that are regarded as default (or standard) complementizers, i.e. Pol. *że*, Russ. *что*, and Slv. *da*.

In our exploratory study we aim to find out where criteria based on word order take us when applying them to randomly selected corpus data. Towards that aim, we want to

- 1) test out whether a rigid bottom-up analysis based on assumptions about the relative position between elements and the occurrence of certain types of units in subsequent clauses yields any conceivable patterns;
- 2) show how far these assumptions as such are testable and ask which consequences our insights may have for syntactic research, primarily into clause-combining from a diachronic perspective.

For the time being, we leave out the question of possible categorizations of these units, and of the respective clause types, from either a synchronic or a diachronic point of view. We also remain agnostic as for whether, and to what extent, patterns of linear positions yield any conclusive evidence concerning main vs embedded contrasts.

We proceed as follows: In Section 2 we introduce the 'syntactic oxymoron' of directive subordination, discuss the challenges presented by the aforementioned directive-optative expressions and develop our working hypotheses. Section 3 introduces our dataset; our exploratory analysis and its results are described in Section 4. We close with a discussion of the findings and their implications against a larger background (Section 5).

## 2. Directive subordination

The uninflected units Pol. *niech*, Russ. *pust'*, Slv. *naj* originate in verbs meaning 'let' (Russ. *pustit'*, \**nehati* > Slv. *naj*, Pol. *niech(aj)*). Together with verbs in the indicative present they serve to mark directive, optative, or permissive speech acts (for other functions see note 2); we will therefore call them DIR-units (or simply DIR), as opposed to elements regarded as standard complementizers (COMP), i.e. Pol. *że*, Russ. *что*, Slv. *da*. Note that we employ DIR and COMP only as labels of surface units without any commitment concerning their functional (semantic) or syntactic qualification (e.g. in terms of auxiliaries, complementizers or other kinds of connectives).

We first present crucial usage types of DIR-units and point out the problems raised by their linguistic assessment (§2.1), before we discuss in which regard, and to which extent, subordination might be compatible with directive speech acts (§2.2). This brings us to the working hypotheses and assumptions (§2.3) which guide us in our corpus-driven analysis of DIR- and COMP-units (§3).

### 2.1. Environments of DIR

DIR can occupy any position from the leftmost periphery of the clause up to immediate adjacency left to the finite verb; see (1-3) for non-initial and (4-6) for initial position. In cases where the finite verb finds itself in the second position, clause-initial and verb-adjacent position coincide (see (5)). Moreover, DIR-units may occur adjacent to preceding COMP-units (see (7-9)). In Polish and Russian, COMP and DIR may be separated by one or more intervening words (see (11-12)), whereas in Slovene such occurrences are extremely rare (see (10)).

DIR non-initial

- (1) Slovene (South Slavic)

*Tu notri stoji zapisano:*

<i>Vsak</i>	<i>naj</i>	<i>vza-me</i>	<i>svoj križ na rame.</i>
each.NOM.SG.M	DIR	take[PFV]-PRS.1PL	

'Here it is written: *Let* each one take his cross on his shoulder.' ('... Everybody may take ...') (IMP)

- (2) Polish (West Slavic)

*Ja tam wierzę swoім metodom,*

<i>a komputerami</i>	<i>niech</i>	<i>się</i>	<i>zajmuj-q</i>	<i>geniusz-e.</i>
	DIR	REFL	deal.with[IPFV]-PRS.3PL	genius-NOM.PL

'I believe in my methods; as for computers, *may* geniuses deal with them.' (PNC)

*Capturing an oxymoron in the wild: Directive subordination in Slavic*

- (3) Russian (East Slavic)  
 (...) *lučše nalogi potratit' na obrazovanie, medicinu, kul'turu,*  
*a ličnye uvlečenija pust' graždan-e oplačivaj-ut iz svoego karmana.*  
 DIR citizen-NOM.PL pay[IPFV]-PRS.3PL  
 'it is better to spend taxes on education, medicine, culture; as for personal hobbies, *may*  
 citizens pay (for them) out of their own pocket.'  
 (RNC)

DIR initial

- (4) Slovene  
*Popoldne je šef<sub>x</sub> končno*  
*ukaza-l, naj me<sub>y</sub> pokliče-jo k njemu<sub>x</sub>.*  
 order[PFV]-PST.(3SG.M) DIR 1SG.ACC call[IPFV]-PRS.3PL to 3SG.M.DAT  
 'In the afternoon the boss<sub>x</sub> finally ordered *that* they call me<sub>y</sub> to (come to) him<sub>x</sub>.'  
 ('... *May* they call me...')  
 (from Uhlik 2018: 412)
- (5) Polish  
*W końcu, żeby zrobić mu przyjemność,*  
*powiedzia-ł-em, niech przyniesi-e.*  
 say[PFV]-PST-1SG.M DIR bring[PFV]-PRS.3SG  
 'In the end, to please him, I said, *may* he bring it.'  
 (PNC)
- (6) Russian  
*No prežde*  
*trebuj-u, pust' začiniščik-i predstan-ut*  
 demand[IPFV]-PRS.1SG DIR instigator-NOM.PL appear[PFV]-PRS.3PL  
*pered zakonom i sudom za vojnu v Čečne i vse soveršennye imi zlodejaniya.*  
 'But first I demand *that* [lit. *let*] the instigators be brought before the law and the court  
 for the war in Chechnya and all the atrocities they have committed.'  
 (RNC)

COMP-DIR, initial (adjacent)

- (7) Slovene  
*Rekel<sub>x</sub> mi<sub>y</sub> je*  
*da naj vas pričaka-m<sub>y</sub>.*  
 COMP DIR 2PL.ACC pick.up[PFV]-PRS.1SG  
 'He<sub>x</sub> told me<sub>y</sub> *that* (I<sub>y</sub>) *should* pick you up.'  
 (<opus.nlpl.eu>)
- (8) Polish  
*Stary odpowiedział,*  
*że niech nawet w więzieniu zgnij-e.*  
 COMP DIR rot[PFV]-PRS.3SG  
 'The old man replied *that* (he) *may* even rot in prison.'  
 (PNC)

(9) Russian

*On<sub>x</sub> govoril ej<sub>y</sub>, čtoby ne vydumyvala<sub>y</sub>,  
 čto pust' pol'zu-et-sja<sub>y</sub> vsem, čem nado.  
 COMP DIR use[IPFV]-PRS.3SG-REFL  
 'He<sub>x</sub> told her<sub>y</sub> not to invent, that (she<sub>y</sub>) may use everything needed.'  
 (RNC)*

COMP + DIR, non-adjacent

(10) Slovene (spoken)

*na pregleda /// tako je /// eee tako rekoč prosili so<sub>x</sub> jih<sub>y</sub>  
 da vendarle nej pridej-o<sub>y</sub>  
 COMP however DIR come[PFV]-PRS.3PL  
 'for inspections /// that's right /// hey, they<sub>x</sub> practically asked them<sub>y</sub> to come anyway'  
 (more lit. '... that however may they<sub>y</sub> come ...')  
 (GOS)*

(11) Polish

*Tusk powiedział,  
 że polityc-y najlepiej niech wróc-q  
 COMP politician-NOM.PL DIR return[PFV]-PRS.3PL  
 do stołu rozmów po wyborach.  
 'Tusk said that it is best for politicians (that) they may return to the negotiating table  
 after the elections.'  
 (PNC)*

(12) Russian

*Armen (...) teper' xodil v osnovnom po prefekturam, polagaja,  
 čto ežednevnoj rabotoj pust' zanimaj-ut-sja drug-ie.  
 COMP DIR deal.with[IPFV]-PRS.3PL-REFL other-NOM.PL  
 'Armen (...) now walked mainly around the prefectures, believing that others should/may  
 do the day-to-day work.'  
 (RNC)*

In all these cases, DIR-units can be regarded as illocutionary force indicating devices (IFIDs),<sup>2</sup> in the sense of Searle (1969). Their syntactic status is less clear. Treating constructions of DIR + V<sub>IND.PRS</sub> as an 'analytical mood', in particular a 'third-person imperative', would make DIR an auxiliary (e.g. Xrakovskij 1992 [2001], for a survey cf. Wiemer & Fortuin *forthcoming*, and Wiemer 2023c: §4.3). However, DIR-units often occur clause-initially (see (4-6)), and many clauses introduced by DIR closely follow another clause that contains an expression able to induce a clausal argument with an implied directive, optative or permissive illocution; compare Slv. *ukazal* '(he) ordered' in (4), Pol. *powiedziałem* '(I) said' in (5), and Russ. *trebuj* '(I) demand' in (6). Under these conditions DIR-units may behave like (emergent) complementizers, i.e. "conjunctions that have the function of identifying clauses as complements" (Kehayov & Boye 2016: 1). Admittedly, this argumentation is expounded

rarely, at least for Russian and Polish; as for Slv. *naj* cf. Uhlik (2018: 411-413) and Sonnenhauser (2021). For references and a critical assessment cf. Wiemer (2021: 84-91; 2023a: 192-205; 2023b: 8-14; 2023c: §§4.1-4.2).

While adjacent combinations of COMP-DIR (see (7-9)) might be rare in corpora oriented toward standard languages, they are by no means exceptional. Since DIR induces directive illocutionary force and COMP is considered to mark a clause as a complement of some (usually preceding) complement-taking predicate (CTP), COMP-DIR combinations violate a presumed hallmark of subordination, namely the loss of independent illocutionary force (see §2.2). A very similar issue arises if DIR-units are analyzed as complementizers (see (4-6)): if their original directive-optative force is retained, we have to acknowledge embedded clauses with independent illocutions. An alternative would be to analyze DIR and COMP-DIR-clauses as quotes, but then COMP can hardly be classified as a complementizer; instead, it would have to be considered a means that sets off clauses as units with independent illocutions and information structure (Wiemer 2023a: 232-240). This looks like another stipulation (alternative to stipulating DIR as a complementizer). In a similar vein, we could treat DIR-clauses as asyndetic complements of anchoring elements in the preceding conjunct,<sup>3</sup> with DIR itself retaining its illocutionary function. This stipulation, again, would end up in admitting independent illocutions in embedded clauses.

## *2.2. Directive speech acts and embedded clauses*

It is often assumed that embedded clauses are void of an independent illocution; at least they are very reluctant to include signs of illocutionary force. For instance, Lehmann (1988) identifies lack of, or constraints on, illocutionary force as the first property that appears on a scale of desententialization (which correlates with subordination) and concludes that “a subordinate clause may not normally have its own illocutionary force” (1988: 193).<sup>4</sup> Similarly, Verstraete (2007: 157-159, 284) *de facto* treats the lack of an independent illocution (‘speech functional value’) as the main (or overarching) property of subordination. See also Cristofaro (2003: 18 and *passim*) from a functional-cognitive point of view as well as Nordström (2010) and Zimmermann (2015: 580) from the position of formal syntax.

The tenet that subordination, in particular embedding, bars independent illocutionary force has been challenged by Krifka (2014; 2023). He demonstrates that modifiers operating on the illocutionary level can show up in clauses that otherwise bear signs of embedding, the most

important one being the presence of an element that looks like a presumably unequivocal complementizer. Compare example (13), taken from a German internet source (Krifka 2023: 160):

(13) German

*da hat er gesagt, dass er offen gestanden keine Ahnung hat, weil du und der Junge euch ständig irgendwie zurückzieht.*

‘then he suddenly said that he *frankly speaking* does not have a clue because you and the boy keep hiding somewhere’

The adverbial *offen gestanden* ‘frankly speaking’ occurs inside a complement clause (introduced by the complementizer *dass* ‘that’), even though it belongs to the reported speaker, not the speaker who uttered example (13). Note, however, that person-deictic expressions (2<sup>nd</sup> person: *du und der Junge*; *euch zurückzieht*) are used from the perspective of the reporting speaker. Shifts of person deixis remain unaccounted for in Krifka’s analysis (see below).

Since we cannot delve into the epistemological background of Krifka’s theory<sup>5</sup> and its details we restrict ourselves to a few pertinent remarks. Krifka’s analysis is only concerned with assertive and interrogative speech acts and the propositions which they convey. Moreover, Krifka heavily relies on word order patterns of English and German, and he takes the appearance of Engl. *that*, Germ. *dass* as a sufficient indication of embedding, i.e. these clause-initial units are understood as flags of clausal arguments regardless of what follows them. However, the clause-initial employment of DIR-units as illustrated in §2.1 is not fully covered by Krifka’s analysis. First, this employment is predominantly relevant for speech acts that are void of propositions, since DIR-units primarily mark directive illocutions (see however note 2). Second, as already emphasized, Slavic languages do not use word order patterns to distinguish main from embedded clauses. Therefore, if clause-initial DIR-units are not preceded by a dedicated complementizer (see (4-6)), shifts of person-deictic expressions remain as the only possible clues of embedding. Unfortunately, often both the presumable reported and the reporting speech act exclusively contain 3<sup>rd</sup> person expressions, so that this potential clue is of no avail. However, whenever original and reporting speech act differ in this respect, person-deictic expressions are shifted in correspondence with the reporting speech act; this happens very systematically (cf. Wiemer 2023b; 2023c). The same seems to be the case with COMP-DIR combinations (see (7-12)), for which different stipulations have been formulated in §2.1. An additional problem with these structures is that we need to be sure that COMP really functions as a complementizer, in the first place (see below).



Let us first examine cases in which only initial DIR introduces a clause. Such clauses happen to occur right after expressions that are suitable CTPs. In these cases, they might be analyzed as direct speech loosely attached to those potential CTPs. Since direct speech preserves the original illocution (Clark & Gerrig 1990, Dirscherl & Pafel 2015, among many others), it seems possible to characterize clauses with initial DIR as quotes that occupy an argument slot of a predicate denoting a certain speech act. Krifka considers analogous cases with assertive speech acts, e.g., by contrasting three ways of conveying other people's utterances, as in (14a-c) cited from Krifka (2014: 77):

(14a) *John said to Mary "I admire Sue".*

(14b) *John told Mary that he admired Sue.*

(14c) *John told Mary he admires Sue.*<sup>6</sup>

The clause *he admires Sue* in (14c) can be understood as the realization of the propositional argument of *told* attached to it asyndetically. Subordination is supported by a person-deictic shift (1<sup>st</sup> > 3<sup>rd</sup> person for *John*) to correspond with the reporting speaker's perspective. This type of clause combining is comparable to clauses with initial DIR-units following clauses with predicates denoting speech acts, except that we are dealing with non-assertive illocutions. Krifka treats the second clause in both (14b) and (14c) as subcategorized by *told*, whereas in (14a) the speech act conveyed by the direct quote is just 'identified' and *said* only denotes the utterance type. The difference between (14b) and (14c) is even subtler: in (14b), "the verb *tell* expresses that an illocutionary act of the type of assertion happens", whereas in (14c) "the verb *tell* does not denote such a speech act, but subcategorizes for this kind of speech act as its argument" (2014: 79). Therefore, since the difference between (14b) and (14c) only consists in the presence vs absence of a complementizer,<sup>7</sup> Krifka's subtle distinction amounts to saying that the complementizer serves as kind of pointer to an imminent illocution.<sup>8</sup> But, first, the complementizer itself does not say much about the TYPE of illocution, and second, when it is lacking (as in (14c)) the only indication of a complementing relation is the semantics of the preceding verb (*tell*, *say*).

Consequently, once the only (?) role of *that* is to point forward to an utterance with its respective illocution, this presumable complementizer might be assigned a broader (and, thus, vaguer) function, which consists in just setting off a subsequent utterance, while the conceptual link between a pair of adjacent clauses arises from a semantic entailment (*say* ⊨ propositional argument) of the speech act verb in the first

clause of that pair. From this angle, COMP-DIR-combinations as in (7-9) above appear in a different light: they may be analyzed as transparent juxtapositions of a ‘pointer’ (with maximal neutrality as for illocutionary force) and an IFID of directive (optative, permissive) speech acts.

Remarkably, this analysis proves adequate also for cases in which there is no potential CTP for COMP to serve as a flag of a clausal argument. See the following Polish example: *zatrzymać* ‘to stop (sb, sth), to detain’ does not entail any clausal argument; instead, *że* helps to continue a narration by introducing a motivation for the activity mentioned in the preceding clause, while *niech poczeka* ‘may he wait’ refers to a directive speech act of the original speaker (*matka* ‘mother’) from the perspective of the reporting speaker.<sup>9</sup>

(15) Polish

*Matka próbowała go zatrzymać, że niech poczeka, rozścieli mu łóżko (...).*

‘Mother tried to stop him, *that may he wait*, she will make his bed for him (...).’

(PNC; W. Myśliwski: *Traktat o łuskaniu fasoli*, 2007; from Wiemer 2023a: 238)

This reasoning, essentially, transfers to cases in which only DIR occurs clause-initially. The only difference is that there is no ‘pointer’, and DIR’s clause-initial occurrence provides a context where DIR might eventually be reanalyzed as a complementizer (which preserves directive illocutionary force). Compare example (16) from Russian, which also shows a 2<sup>nd</sup> > 3<sup>rd</sup> person shift, since it reflects the perspective of the speaker who anticipates a demanded speech act:<sup>10</sup>

(16) Russian

*I skazite svoim doverennym – pust’ ne boltajut.*

‘And *tell* your trusted ones – *may* they not chat.’

(RNC; A. Lazarčuk & M. Uspenskij: *Posmotri v glaza čudovišč*, 1958; from Wiemer 2023b: 17)

Therefore, a joint consideration of our and Krifka’s analyses suggests that person-deictic expressions adapt to the reporting speaker ‘earlier’, or more readily, than the illocution of the reported speech act. This is why utterances like (15-16), but also (14c), defy a clear analysis in terms of either direct or indirect speech – with the consequence that also the embedding of illocutions proves to be a gradable phenomenon. In this perspective, symptoms of independent illocutions are to be expected in quote-like clauses that are also embedded to some extent. Apart from an account of person-deictic shifts, our approach also differs from Krifka’s in that we do not accept *prima facie* COMPs (like Germ. *dass*, Engl. *that*) to always function as flags of clausal arguments, even if the conditions are favorable.<sup>11</sup> Their syntactic function may be entangled with (or even overridden by) functions of marking discourse continu-

ity (see example (15)) or (self)quotes. The latter differ from the former mainly in that they lack person-deictic shifts.<sup>12</sup>

Finally, even though Krifka (and others) have shown that ‘root phenomena’ like illocutionary force do occur in natural settings of embedding, in natural discourse such phenomena are difficult to recognize. That is, modifiers operating on an illocutionary level (as in example (13)) are comparatively rare in corpus and other data. We thus assume that usually speakers are simply not confronted with mismatches between, e.g., person-deictic expressions and illocutionary markers in clauses that might be considered as structurally dependent on (or, in particular, embedded in) other clauses or predicates; and if they are this does not create problems for ‘bringing one’s message across’.

The analytical problems connected to all imaginable stipulations pointed out in §2.1 make DIR-clauses and, even more so, COMP-DIR-clauses look like syntactic oxymora (if subordination is understood in a categorial way) or as hybrids (if subordination is treated as a gradable phenomenon). Because of this conundrum, we will refer to COMP-DIR combinations and to DIR-clauses as ‘directive subordination’.

### *2.3. Assumptions and working hypotheses*

In view of the problems outlined above, the challenge consists in identifying patterns without having to categorize elements (DIR, COMP) whose status is not (yet) clear. We thus approach directive subordination and its status within clause-combining by identifying the clause-internal and cross-clausal syntagmatic behavior of DIR-units compared to the behavior of units that are usually identified as canonical (or entrenched) complementizers. For this purpose, we explore the distribution of DIR-units and use it as a diagnostic tool to search for links between word order preferences and clause types in a strictly corpus-driven and maximally surface-oriented approach. Although we acknowledge hierarchical orderings among a (verbal) predicate (V) and its arguments (S, O, etc.), we do not imply the existence of sequences other than directly observable ones. Clause-internally, we look for word order patterns in clauses containing COMP, DIR or COMP-DIR, i.e. adjacent and non-adjacent co-occurrence of COMP and DIR within the same clause. With respect to cross-clausal patterns, we look for potential CTPs by considering all expressions that may serve as a possible anchor of a subsequent clause and thus foster the emergence of cross-clausal relations (see §3.2). That is, on the one hand, we expand the concept of word order to relations between function words (connectives) and finite verb. On the other hand, our agnostic perspective on the anchoring patterns

allows for insights into mechanisms of clausal complementation beyond presumed, ‘classical’ sets of CTPs. This avoids reasoning that runs the risk of becoming circular (‘this is complementation because we have a CTP’).

Thus, by comparing clauses containing COMP and/or DIR, we explore the possibility of identifying clause-internal and cross-clausal patterns, while we remain agnostic as for main vs embedded contrasts. We only use clause-initial COMP and DIR as proxies for subordination, because first and foremost we are interested in whether DIR- and COMP-DIR structures differ systematically from COMP-structures, and whether their patterns change over time, in particular whether structures containing DIR get closer to COMP-structures.

The assumption that auxiliaries tend to be more tightly bound (in morphosyntactic terms) to ‘their’ lexical verbs, is problematic especially w.r.t. Slavic languages (cf. Wiemer 2023a, Wiemer & Fortuin *forthcoming*). However, for operational purposes, we consider the clause-initial and non-adjacent position (relative to the predicate) indicative for a connective to be leaning toward the complementizer pole, while a stronger tendency of being close, or even attached, to the verb could be seen as indicative of auxiliaries.

From the observations discussed for examples (1-12) and the general assumptions on the relevance assigned to connectives and to word order for subordination, we derive six hypotheses (H1-H6), which result from a decomposition of these concepts as far as they concern the linear distribution of clausal elements.

- [H1] Provided canonical complementizers occupy clause-initial position, they should allow for a larger distance to the finite verb than elements that primarily modify the finite verb and tend to occupy a position adjacent to it. Thus, there should be an observable difference in terms of distance to the finite verb between more canonical complementizers (COMP) and less canonical ones (DIR).
- [H2] If there is no difference in word order for main and subordinate clauses in Slavic, we should not find any difference concerning the position of elements in COMP- and DIR-clauses.
- [H3] Clauses with adjacent COMP and DIR show more variation with respect to the position of the finite verb than either of them by itself, because their positions are influenced by two different principles (see H1).
- [H4] If patterns change over time, the patterns for DIR and COMP-DIR converge with the patterns for COMP (because they start behaving like canonical complementizers).

[H5] From a synchronic perspective, there are frequency biases in the occurrences of COMP, COMP-DIR and clause-initial DIR in the immediate vicinity of potential CTPs.

[H6] In diachronic comparison, there are frequency shifts for identical syntactic types of potential CTPs (regarding the same language).

### *3. Data: samples and annotation*

We searched for the relevant elements in corpora of Russian, Polish and Slovene (see References), both separately and in combination (with a maximal distance of three word forms between COMP and DIR). To include a diachronic perspective, we considered data from two temporal layers, viz. ca. 17<sup>th</sup>-19<sup>th</sup> ('old') and ca. 1980-2021 ('cont(emporary)'). These two layers roughly capture two sociolinguistic stages in the history of the three languages: the period stretching from an increasing literary production in the respective vernaculars up to the first attempts of standardization, and the contemporary standard languages. This sociolinguistic characterization of our samples ensures comparability also for the 'old' samples, for which the corpora differ in terms of actual temporal coverage for the three languages.

We randomized the search results and filtered them according to the following criteria:

- General: We excluded structures in which neither COMP nor DIR appear at the left edge of a clausal unit.
- DIR: We excluded cases where DIR does not signal directive-optative illocutionary force, i.e. cases with non-curative, conditional or concessive functions (see note 2).

Our final dataset includes samples with 50 random tokens of COMP, DIR and COMP-DIR structures per language and period. An exception is the sample Pol\_old for COMP-DIR since the search in the KorBa corpus resulted in only 13 instances. An additional search in two other Polish historical corpora (see References) drew another 12 examples, yielding a total of 25 instances.

Central for our concerns is the position of the finite verb in relation to the connective. We approach this issue via word order patterns, focusing on position and on distance. Concerning position, we concentrate on the order of the finite verb (Vfin) and its main arguments and on the position of Vfin within the clause, taking the connective as point of reference (initial: I, preceded by a clitic element: C, medial: M, final: F). Concerning distance, we measure the distance between Vfin and

connective and, for the COMP-DIR samples, the distance between both connectives. Distance is measured in terms of syntactic constituents as well as syllables. In order to gain insight into the anchoring of clauses containing COMP and/or DIR with a preceding or following clause, i.e. cross-clausal patterns, we annotated the presence of a potential predicative head (also known as CTP) to the left (immediately or further away) or to the right of the connective and its part of speech. Our annotation also accounted for a rough lexical classification of such potential CTPs, but we will not deal with this parameter in this article.

For detailed information on the data basis and the annotation, see Supplementary Information I ‘Data Preparation’, which is stored with a CC-license under <[github.com/IkerSalaberri/Supplementary-materials-for-Special-Issue-Comparative-approaches-to-the-DB-of-clause-types-.git](https://github.com/IkerSalaberri/Supplementary-materials-for-Special-Issue-Comparative-approaches-to-the-DB-of-clause-types-.git)> (see also the introductory article to this issue).

#### *4. Analysis and results*

There are certain properties in the data set that might skew the results and that have to be accounted for when interpreting them. The sample Pol\_old (17<sup>th</sup>-19<sup>th</sup> centuries) contains only 25 non-random tokens for COMP-DIR (see §3). The Russian samples show a considerable number of tokens without a finite verb (annotated ‘dna’). This is not the case for the Slovene and Polish samples. Moreover, we took the following decisions to make the data manageable and interpretable. Vpos=M (M = medial) also includes the ‘almost last’ positions (second to last, etc.), whereas Vpos=F (F=final) means only the very last position. Thus, diachronic changes in the Vpos=M position (e.g. in Russian and Polish) might simply result from an increasing length of utterances. In order to downsize the considerable variation in the linear occurrence of arguments, their surface forms (NPs, (clitic) pronouns and clauses) have been merged. One should keep in mind, however, that their position might have been influenced by their varying weight (Hawkins 1994; 2004, Kizach 2012) in terms of morphemes or syllables. The DIR-tokens were syntactically and semantically preselected (see §3). Thus, the fact that we filtered out cases that did not show DIR at the left edge of the clause might have biased the sampling of data points regarding its position relative to Vfin.

In the following, we discuss the data in regard to the hypotheses formulated in §2.2. The numbers of figures refer to the Supplementary Information II ‘Exploratory analysis’, which is stored with a CC-license under <[github.com/IkerSalaberri/Supplementary-materials-for-Special-Issue-Comparative-approaches-to-the-DB-of-clause-types-.git](https://github.com/IkerSalaberri/Supplementary-materials-for-Special-Issue-Comparative-approaches-to-the-DB-of-clause-types-.git)> (see also

the introductory article to this issue). For convenience, the hypotheses are fully repeated.

[H1] Provided canonical complementizers occupy clause-initial position, they should allow for a larger distance to the finite verb than elements that primarily modify the finite verb and tend to occupy a position adjacent to it. Thus, there should be an observable difference in terms of distance to the finite verb between more canonical complementizers (COMP) and less canonical ones (DIR).

H1 finds partial support in the data. As shown in Fig. 1, in Russian and Polish, *Vfin* is generally placed later in the clause (i.e. *Vpos*=M/F) for COMP than for DIR and partly for COMP-DIR. The latter two have a stronger tendency for *Vfin* being placed in initial or middle position (i.e. *Vpos*=I/M). When comparing the distribution across the three languages, we find that DIR is more frequent with final and middle positions of *Vfin* in Polish, while in Russian (and Slovene, but see below) the initial and middle positions are more preferred; the same holds true for COMP-DIR.

This observation correlates with differences concerning the distance between COMP and DIR: in *Pol\_old* 7 out of 25, and in *Pol\_cont* 22 out of 50 cases have these connectives in non-adjacent position, while Slovene only has 3 (out of 50) in *Slv\_old* and none in *Slv\_cont*. Russian falls in-between: 8 cases in the old and 10 in the contemporary stage (for each out of 50) are non-adjacent. Thus, if left-edge position and linear association with COMP are considered indicative of the status of DIR, Polish shows the least tendency for DIR to behave like a complementizer at either stage. The distance between COMP and DIR in Polish also supports the tendency for *Vfin* to appear farther away from COMP.

Moreover, a look at the syllables intervening between the (last) connective and *Vfin* (see Fig. 2-3) reveals that there are cases where *Vpos*=M has more intervening material than *Vpos*=F. This indicates that the number of syllables should also be taken into account and might be more revealing than a focus on the position of *Vfin*. COMP and *Vfin* are separated by more syllables than DIR and COMP-DIR in Russian and Polish; this holds true especially for the contemporary varieties.

Slovene differs from Polish and Russian. It shows a more compact behavior, i.e. there are less differences between COMP, COMP-DIR and DIR (Fig. 5), inasmuch as *Vfin* tends to follow any of them very closely (*Vpos*=CI; Fig. 1 and 4).

[H2]If there is no difference in word order for main and subordinate clauses in Slavic, we should not find any difference concerning the position of elements in COMP- and DIR-clauses.

Apart from the absence of a clear pattern concerning the position of the finite verb and its distance from the connective (see findings for H1), all three languages show very different distributions. However, for each language the frequency patterns of the connectives between periods more or less remained the same.

As for the ordering of arguments, i.e., nominal expressions filling valency slots, and the finite verb (variable *Vfin\_arg*), we could not identify clear patterns either. Our annotation resulted in about 160 different options. We thus decided to simplify the annotation by collapsing all types of arguments (i.e. with and without p/a/c-index, see Supplementary Information ‘Data Preparation’). This resulted in slightly less than 100 ordering options. From those, we considered the most frequent patterns, i.e. <V, DO, S>, <V, DO>, <V, S> and <V> and divided them into patterns with and without S (Fig. 6-8). This revealed the following tendencies: For the patterns with S, i.e. <V, DO, S> and <V, S>, we observe a less pronounced peak for patterns with S preceding V and DO for DIR than for COMP and even more so for COMP-DIR (Fig. 6). For patterns without S, there is a slightly stronger preference for DIR to have V preceding DO than for COMP in all three languages; the preference for V\_DO does not change across time. COMP is a bit more balanced in these respects (Fig. 7-8).

[H3]Clauses with adjacent COMP and DIR show more variation with respect to the position of the finite verb than either of them by itself, because their positions are influenced by two different principles (see H1).

This hypothesis could not be confirmed. Fig. 9 suggests that COMP in Russian and Polish has more dispersion than COMP-DIR and DIR. It therefore seems that for these two language, COMP-DIR and DIR have more commonalities with each other than COMP with either of the other two, while for Slovene there are no obvious differences.

[H4]If patterns change over time, the patterns for DIR and COMP-DIR converge with the patterns for COMP (because they start behaving like canonical complementizers).

Our data does not support this hypothesis directly. However, there are diachronic changes with respect to the distance between con-



nective and Vfin. In Polish, we observe a slight shift from Vpos=F to Vpos=M with COMP and COMP-DIR. At first glance, this contradicts H4. However, if we look at the distance between the respective connective and Vfin measured in syllables, we notice an increase in syllables for Vpos=M, particularly with COMP. This is also true for Russian. At the same time, the number of syllables between the connective and the finite verb decreases for Vpos=F (Fig. 10). Apparently, there is a diachronic tendency to increase the distance between the connective and Vfin in Russian and Polish, which could be interpreted as a partial support for the hypothesis. These findings suggest that the parameter ‘distance’ needs to be split into two independent parameters (position of Vfin within the clause and linguistic material intervening between the connective and Vfin) also for a diachronic analysis (see H1 above).

Slovene does not confirm H4. The value Vpos=M increases for all connectives. In addition, Vpos=F decreases with COMP-DIR – but again, there is no obvious change in terms of syllables between connective and finite verb.

[H5]From a synchronic perspective, there are frequency biases in the occurrences of COMP, COMP-DIR and clause-initial DIR in the immediate vicinity of potential CTPs.

This hypothesis could be confirmed. Most cases with no suitable CTP in the vicinity (annotation: dna) occur with DIR, whereas CTP=dna is rare with COMP and COMP-DIR. This is true for Russian and Polish. In Russian almost all ‘dna’-values occur with DIR (*pust’*). Slovene has fewer cases with no suitable CTP for DIR, and, in addition, DIR is not very different from COMP and COMP-DIR (Fig. 11 and 12).

Unsurprisingly, CTPs are predominantly verbs (CTP\_pos=v(erb)). DIR shows the lowest degree of variation with respect to CTP\_pos, the CTPs are mostly verbs. This is particularly true for Russian and Polish. This last fact should be treated with caution since DIR-units often do not have a suitable CTP in their vicinity whereas with COM and COMP-DIR the presence of a CTP is the rule.

[H6]In diachronic comparison, there are frequency shifts for identical syntactic types of potential CTPs (regarding the same language).

In our data, there is no discernible pattern that would support this hypothesis. There are single trends, but it is yet unclear how they fit into what kind of overall picture. This is partly due to the fact that

cases with CTP\_pos = dem(onstratives), CTP\_pos = pred(icatives) or CTP\_pos = qu(antifiers) are infrequent.

In Russian, CTP\_pos = pred slightly decreased, in particular with COMP, less so with COMP-DIR. The number of CTP\_pos = dem for COMP and COMP-DIR slightly increased. For Polish we observe an increase of CTP\_pos = n(oun) with COMP, less so with COMP-DIR. Cases with CTP\_pos = dem for COMP and CTP\_pos = qu for COMP-DIR decreased. In Slovene, there is a general decrease of cases with no suitable CTP in the vicinity of the three connectives (CTP\_pos = dna) (Fig. 12).

In conclusion, we can formulate three main results:

- (i) Slovene differs from Russian and Polish w.r.t. almost every of the parameters investigated here.
- (ii) As for the right context, i.e. clause internally, the behavior of COMP markedly differs from both COMP-DIR and DIR (i.e. COMP-DIR rather behaves like DIR, not like COMP).
- (iii) With regard to the left context, i.e. across clauses, there is a divide between DIR, on the one hand, and COMP-DIR and COMP, on the other, caused by the frequent absence of a suitable CTP with DIR.

## *5. Discussion and outlook*

Focusing on Russian, Polish and Slovene as representatives of the three major Slavic branches, we operationalized the correlation assumed in the literature between word order and clause type by taking clause-initial function words as indicators of (some degree of) subordination. In line with traditional assumptions, we considered COMP-elements as indicators of complement clauses. This allowed us to look for similarities and differences in comparison to clauses displaying DIR-elements, including COMP-clauses with a DIR-element. On this basis, we assessed the correlation of word order and clause type (defined via connectives) in a synchronic and diachronic perspective by testing six hypotheses.

Our exploratory analysis of the data does not supply any straightforward responses to our hypotheses. What are the reasons for this apparently unsatisfying result? Have we used insufficient data or bad concepts? What is to blame – the data, our analysis, or the concepts?

We took the corpus data at face value, i.e. as good data in a qualitative sense, notwithstanding possible deviations from standard varieties. Linguistic description and meaningful analysis must be able to cope with data occurring ‘in the wild’. From a quantitative perspective, our samples might be too small to detect relevant patterns concerning potential

correlations of the various implementations of word order and the connective elements. It might also be the case that the periods covered by the labels ‘old’ and ‘cont(emporary)’ are too large, or that the distance between them is too small to detect any developments. There is also the possibility that a more sophisticated statistical analysis might have revealed particular correlations that otherwise escape the explorers’ eyes. Moreover, the clause level might not be the right place at all to detect word order changes, and changes rather occur between pairs of constituents (as suggested by Jing *et al.* 2023).

Last, but not least, we cannot exclude that it was our annotations that failed to detect patterns. The rationale underlying each variable was derived from the assumptions as formulated in the literature, but their empirical implementation turned out less straightforward than their draft on the conceptual easel. There are several options of how to conceptualize word order, all of which are problematic when dealing with languages that often drop their subjects, such as Polish and Slovene, or the copula verb, such as Russian. Moreover, there are several perspectives on the position of the finite verb (if available).<sup>13</sup> It seems that ‘word order’ covers quite different phenomena, which also comprises the relation between the predicate (and its arguments) and clause connectives as well as linear orderings below clause level. Here we accounted for the former (not the latter), but our data did still not reveal any obvious patterns or correlations, no matter how we looked at positions between elements. This finding corroborates recent insights that prove word order to be a fuzzy and gradable concept (cf. Levshina *et al.* 2023), probably no less than that of subordination (see §2.2). Taken together, the observations related here suggest that, without further qualifications, word order is of little help for establishing clause type distinctions.

In §2.1 we pointed out various stipulations for analyzing DIR and COMP-DIR structures. All of them lead to a dead end, unless we admit that embedded clauses may reveal properties of independent illocutions to some extent, and under certain conditions. These conditions require further exploration. Simultaneously, we are left with the question how complementation and complementizers might be diagnosed. After all, findings derived from our corpus-study and the annotation grid cannot by themselves be considered as proof (or disproof) of either main vs embedded contrasts or of complementizerhood of clause-initial connectives.

This brings us to general methodological issues. From the start, we emphasized that the concept of subordination itself has hardly ever been challenged from an analytical, let alone empirical point of view. In fact, many studies on syntactic change presuppose main vs embedded contrasts (or asymmetries) as clear and settled; a good case in point are

varying, in part contradictory, claims as for whether innovations in morphology are initiated and propagated first in main or embedded clauses or whether, conversely, it is rather main or embedded clauses that retain obsolete morphosyntactic patterns (e.g. Stockwell & Minkova 1991, Bybee 2002, Axel 2007, Ledgeway 2021). While we do not question the existence of asymmetrical structures in clause-combining, and thus subordination as a gradable concept, we are more skeptical about their diagnostics, in particular when it comes to earlier and non-standardized stages of languages.

A way out will be multivariate approaches along the lines suggested in Bickel (2010), that analyze given structures via a set of variables and may thus derive similarity clusters from the data. To make this approach feasible for the kind of corpus-driven bottom-up approach proposed in this paper, with the aim of tracing the dynamics of clause-combining within and across varieties, the variables must be retrievable in corpora. This is possible with the parameters applied here, but still challenging for semantic and pragmatic features, such as person-deictic shifts or the role of illocutionary force (see §2.2).

Finally, a bottom-up approach should also be able to deal with indeterminate cases, i.e. phenomena that cannot be properly handled by applying categorial distinctions which are convenient for linguistic theorizing and the annotation of corpora. As for subordination, indeterminate cases include the continuum between direct and indirect speech or combinations of more than one clause-initial element (see §2.1). Indeterminate cases may not be rare at all, but even if they seem exceptional or non-canonical, they are so primarily for categorial linguistic description and prescriptive grammars based on written standard varieties; speakers, obviously, live very well with this ‘deviance’. This makes such in-between cases relevant in particular for non-standardized varieties. At any rate, they are key in understanding historical processes like time (in)stability and pathways of change.

In sum, we hope to have shown that traditional tools and premises of syntactic analysis do not suffice to meet the requirements needed to describe natural data fully and adequately; often they may even obliterate, or block, the view on data (and linguistic behavior) which deserves to be looked at with less category-framed and/or prescriptive prejudice. In particular, linking the question of word order to that of subordination amounts to juggling with two very complex concepts (each of which needs to be ‘decomposed’), and this makes us lose the safe ground of empirical evidence. To put it differently: connecting two unknowns does not constitute an observation. We should avoid getting trapped in vicious circles.

## Abbreviations

1, 2, 3 = first, second, third person; ACC = accusative; COMP = (standard) complementizer; CTP = complement-taking predicate; DAT = dative; DIR / DIR = directive unit; dna = does not apply; IFID = illocutionary force indicating devices; IND = indicative; IPFV = imperfective; M = masculine; NOM = nominative; PFV = perfective; PL = plural; PRS = present; PST = past; REFL = reflexive; SG = singular. As regards GOS, IMP, PNC, RNC etc., see the list of sources at the end of the bibliographical references.

## Notes

<sup>1</sup> See examples (4,7,9,10) in §2.1. For further discussion see §2.2.

<sup>2</sup> DIR-units also tend to acquire other functions, such as non-curative, conditional or concessive functions (cf. Dobrušina 2019 for Russian). In these functions, DIR units scope over propositions and can occur with finite verbs in a past tense (Wiemer 2024: §3.3). Such DIR-tokens were excluded from the samples (see §3.1).

<sup>3</sup> Such anchoring elements seem to be comparatively rare for DIR-clauses (see on [H5] in §4).

<sup>4</sup> Among exclusions Lehmann adduces non-restrictive relative clauses, which are linked to the head noun (or NP) not via dependency, but via sociation (1988: 194).

<sup>5</sup> Krifka's global aim is the incorporation of illocutionary acts into a semantic theory; this theory should also allow for the inclusion of speech acts into a model of recursion and a dynamic representation of how speech acts contribute to updates of common ground between interlocutors. For this purpose, it has to be shown that illocutions can be treated as abstract objects, so that they become arguments of predicates (i.e. CTPs) and targets of modifiers (e.g. sentence adverb(ial)s); cf. in particular Krifka (2014: 85). In Krifka (2023) this goal is extended to also include discourse moves. Importantly, judgment, commitment and act (i.e. moves in discourse) are different layers that can host different kinds of modifiers and heads (2023: 155).

<sup>6</sup> The German equivalent of the second clause in (14c) would have V2-syntax, apart from possible subjunctive morphology on the verb (Germ. *Konjunktiv I*). Like word order, the latter would be futile for Slavic languages, since they do not employ non-indicative mood to mark reported speech.

<sup>7</sup> In these examples, there is also tense shift, which however is likewise normally absent in Slavic languages (see §1).

<sup>8</sup> Cf. Davidson (1968) for a similar treatment.

<sup>9</sup> In addition, since *niech* practically never occurs with 2<sup>nd</sup> person of the finite verb (in direct speech), we may assume that it 'replaces' an imperative (*poczekaj!* 'wait!') of the imagined reported speech act.

<sup>10</sup> Here reporting and reported speaker collapse since the potential CTP itself marks a directive speech act.

<sup>11</sup> Moreover, Krifka qualifies utterances with parenthetical comments (as in *It's just started to rain, he said*) as instances in which the part commented on by the parenthetical is 'embedded' (2023: 137). This conflates the syntactic notion of embedding (e.g. in Lehmann 1988) with discourse-pragmatic functions of parentheticals, which often fulfill equivalent functions in terms of information structure (Wiemer 2023a: 223-225).

<sup>12</sup> Cf. Letučij (2023) on *čto* 'that' in colloquial Russian, which reveals properties very similar to Pol. *że* and other COMPs (with and without subsequent DIR).

<sup>13</sup> The Russian samples have numerous tokens without a finite verb (see §4).

## Bibliographical References

- Axel, Karin 2007. *Studies on Old High German syntax: left sentence periphery, verb placement and verb-second*. Amsterdam / Philadelphia: Benjamins.
- Axel-Tober, Karin 2012. *(Nicht-)kanonische Nebensätze im Deutschen. Synchrone und diachrone Aspekte*. Berlin / Boston: De Gruyter Mouton.
- Bickel, Balthasar 2010. Capturing particulars and universals in clause linkage. A multivariate analysis. In Bril, Isabelle (ed.), *Clause linking and clause hierarchy. The syntax and pragmatics interface*. Amsterdam / Philadelphia: Benjamins. 51-102.
- Bybee, Joan 2002. Main clauses are innovative, subordinate clauses are conservative: consequences for the nature of constructions. In Bybee, Joan & Noonan, Michael (eds.), *Complex sentences in grammar and discourse: essays in honor of Sandra A. Thompson*. Amsterdam / Philadelphia: Benjamins. 1-18
- Clark, Herbert H. & Gerrig, Richard J. 1990. Quotations as demonstrations. *Language* 66-4. 764-805.
- Cristofaro, Sonia 2003. *Subordination*. Oxford: Oxford University Press.
- Davidson, Donald 1968. On saying that. *Synthese* 19. 130-146.
- Diessel, Holger 2015. Adverbial subordination. In Luraghi, Silvia & Parodi, Claudia (eds.), *The Bloomsbury Companion to Syntax*. London etc.: Bloomsbury. 341-353.
- Dirscherl, Fabian & Pafel, Jürgen 2015. Die vier Arten der Rede- und Gedanken-darstellung. Zwischen Zitieren und Referieren. *Linguistische Berichte* 241. 3-47.
- Dobrušina, Nina 2019. Status konstrukcii s časticami *pust'* i *puskaj* v ruskom jazyke [The status of constructions with the particles *pust'* and *puskaj* in Russian]. *Russian Linguistics* 43. 1-17. <doi.org/10.1007/s11185-018-09208-0>.
- Gast, Volker & Diessel, Holger 2012. The typology of clause linkage. In Gast, Volker & Diessel, Holger (eds.), *Clause linkage in cross-linguistic perspective*. Berlin / Boston: De Gruyter Mouton. 1-36.
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.
- Jing, Yingqi; Widmer, Paul & Bickel Balthasar 2023. Word order evolves at similar rates in main and subordinate clauses: Corpus-based evidence from Indo-European. *Diachronica* 40-4. 532-556. <doi.org/10.1075/dia.20035.jin>.
- Kehayov, Petar & Boye, Kasper 2016. Complementizer semantics – an introduction. In Boye, Kasper & Kehayov, Petar (eds.), *Complementizer semantics in European languages*. Berlin / Boston: De Gruyter Mouton. 1-11.
- Kizach, Johannes 2012. Evidence for weight effects in Russian. *Russian Linguistics* 36. 1-20. <doi:10.1007/s11185-012-9096-0>.
- Kortmann, Bernd 1997. *Adverbial subordination (A typology and history of adverbial subordinators based on European languages)*. Berlin / New York: Mouton de Gruyter.

- Krifka, Manfred 2014. Embedding illocutionary acts. In Roeper, Tom & Speas, Margaret (eds.), *Recursion: Complexity in cognition*. Cham etc.: Springer. 59-87.
- Krifka, Manfred 2023. Layers of assertive clauses: propositions, judgements, commitments, acts. In Hartmann, Jutta M. & Wöllstein, Angelika (eds.), *Propositionale Argumente im Sprachvergleich (Theorie und Empirie)*. Tübingen: Narr Francke Attempto. 115-181.
- Ledgeway, Adam 2021. V2 beyond borders: the *Histoire Ancienne jusqu'à César*. *Journal of Historical Syntax* 29-5. 1-65.
- Lehmann, Christian 1988. Towards a typology of clause linkage. In Haiman, John & Thompson, Sandra A. (eds.), *Clause combining in grammar and discourse*. Amsterdam / Philadelphia: Benjamins. 181-225.
- Letučij, Aleksandr B. 2023. Osobennosti razgovornogo upotreblenija sojuza čto: citativnost' i drugie nestandartnye svojstva [Peculiarities in the colloquial use of the conjunction čto: quotativity and other non-standard properties]. *Voprosy jazykoznanija* 2023,3. 27-59.
- Levshina, Natalia\*; Namboodiripad, Savithry\*; Allasonnière-Tang, Marc; Kramer, Mathew; Talamo, Luigi; Verkerk, Annemarie; Wilmoth, Sasha; Garrido Rodriguez, Gabriela; Gupton, Timothy Michael; Kidd, Evan; Liu, Zoey; Naccarato, Chiara; Nordlinger, Rachel; Panova, Anastasia & Stoyanova, Natalia 2023. Why we need a gradient approach to word order. *Linguistics* (aop). <doi:10.1515/ling-2021-0098>.
- Nordström, Jackie 2010. *Modality and subordinators*. Amsterdam / Philadelphia: Benjamins.
- Raible, Wolfgang 1992. *Junktion. Eine Dimension der Sprache und ihre Realisierungsformen zwischen Aggregation und Integration*. Heidelberg: Winter.
- Searle, John R. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Siewierska, Anna (ed.) 1998. *Constituent order in the languages of Europe*. Berlin / New York: De Gruyter.
- Siewierska, Anna & Uhlířová, Ludmila 1998. An overview of word order in Slavic languages. In Siewierska, Anna (ed.), *Constituent order in the languages of Europe*. Berlin / New York: De Gruyter. 105-150. <doi:10.1515/9783110812206.105>.
- Sonnenhauser, Barbara 2021. Slovene *naj*: An (emerging) clausal complementiser? In Wiemer, Björn & Sonnenhauser, Barbara (eds.), *Clausal complementation in South Slavic*. Berlin / Boston: De Gruyter Mouton. 442-475.
- Stockwell, Robert P. & Minkova, Donka 1991. Subordination and word order change in the history of English. In Kastovsky, Dieter (ed.), *Historical English syntax*. Berlin / New York: De Gruyter. 367-408.
- Uhlik, Mladen 2018. O *naj* in *pust'* v slovensko-ruski sopostavitvi [On *naj* and *pust'* in Slovene and Russian]. *Slavistična revija* 66-4. 403-419.
- Verstraete, Jean-Christophe 2005. Two types of coordination in clause combining. *Lingua* 115. 611-626.
- Verstraete, Jean-Christophe 2007. *Interpersonal grammar and the analysis of adverbial clauses in English*. Berlin / New York: Mouton de Gruyter.
- Weiss, Daniel 1989. Parataxe und Hypotaxe – Versuch einer Skalierung. In Girke, Wolfgang (ed.), *Slavistische Linguistik 1988*. München: Sagner. 287-322.

- Wiemer, Björn 2021. A general template of clausal complementation and its application to South Slavic: theoretical premises, typological background, empirical issues. In Wiemer, Björn & Sonnenhauser, Barbara (eds.), *Clausal complementation in South Slavic*. Berlin / Boston: De Gruyter Mouton. 29-159.
- Wiemer, Björn 2023a. Between analytical mood and clause-initial particles – on the diagnostics of subordination for (emergent) complementizers. *Zeitschrift für Slawistik* 68-2. 187-260. <doi.org/10.1515/slav-2023-0012>.
- Wiemer, Björn 2023b. Directive-optative markers in Slavic: observations on their persistence and change. *Linguistica Brunensia* 71-1. 5-45.
- Wiemer, Björn 2023c. Clause-initial connectives, bound and unbound: Indicators of mood, of subordination, or of something more fundamental? *Slavia Meridionalis* 23 (Special issue: *Comparative and typological approaches to Slavic languages*. Ed. by Jakub Banasiak, Julia Mazurkiewicz-Sułkowska, Bożena Rozwadowska, Dorota Klimek-Jankowska). <doi:10.11649/sm.3194>.
- Wiemer, Björn & Fortuin, Egbert *forthcoming*. Mood. In Greenberg, Marc L. (ed.), *Encyclopedia of Slavic Languages and Linguistics online*. Brill: Leiden. <referenceworks.brillonline.com/browse/encyclopedia-of-slavic-languages-and-linguistics-online>
- Xrakovskij, Viktor S. (ed.) 1992. *Tipologija imperativnyx konstrukcij*. Saint-Petersburg: Nauka. [English translation: 2001. *Typology of Imperative Constructions*. München: Lincom.]
- Zimmermann, Ilse 2015. The Russian subjunctive. In Zybatow, Gerhild; Biskup, Petr; Guhl, Marcel; Hurtig, Claudia; Mueller-Reichau, Olav & Yastrebova, Maria (eds.), *Slavic grammar from a formal perspective. (The 10th Anniversary FDSL Conference, Leipzig 2013)*. Frankfurt am M. etc.: Lang. 579-594.

## Sources

- Beethoven: <diaspol.uw.edu.pl/polniem/#!/>
- GOS: Korpus Govorjene Slovenščine / Corpus of spoken Slovene, <www.korpus-gos.net>
- IMP: IMP language resources for historical Slovene, <nl.ijs.si/imp>
- KorBa: The Electronic Corpus of 17th- and 18th-century Polish Texts (up to 1772), <korba.edu.pl/query\_corpus>
- Korpus XIX wieku: <diaspol.uw.edu.pl/XIX/#!/>
- PNC: Polish National Corpus, <nkjp.pl>
- RNC: Russian National Corpus, <ruscorpora.ru>