



UvA-DARE (Digital Academic Repository)

Temporal 'Since' in Slovak: Conjunction(s) and Aspect Choice

A Corpus Study

Kyselica, P.; Genis, R.M.

DOI

[10.2478/jazcas-2019-0051](https://doi.org/10.2478/jazcas-2019-0051)

Publication date

2019

Document Version

Final published version

Published in

Jazykovedný časopis: vedecký časopis pre otázky teórie jazyka

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Kyselica, P., & Genis, R. M. (2019). Temporal 'Since' in Slovak: Conjunction(s) and Aspect Choice: A Corpus Study. *Jazykovedný časopis: vedecký časopis pre otázky teórie jazyka*, 70(2), 200-215. Article 7. <https://doi.org/10.2478/jazcas-2019-0051>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

TEMPORAL ‘SINCE’ IN SLOVAK: CONJUNCTION(S) AND ASPECT CHOICE – A CORPUS STUDY

PAULA KYSELICA – RENÉ M. GENIS
University of Amsterdam, The Netherlands

KYSELICA, Paula – GENIS, M. René: Temporal ‘since’ in Slovak: conjunction(s) and aspect choice – a corpus study. *Journal of Linguistics*, 2019, Vol. 70, No 2, pp. 200 – 215.

Abstract: It has recently been shown by especially [1] through [4] and [12] for Russian and by [8] and [9] for Polish that conjunctions corresponding to Dutch *sinds* (cf. also [1], [2], [3]) and English *since* (cf. also [7], [10]) have temporal functions, which are subject to restrictions on the choice of tense and aspect. Ultimately these restrictions can be related to the semantic input of tense and aspect into complex sentences with these connective items. For Polish extensive data provided by corpus research enabled us to shed light on the usage and restrictions in this area and also to establish which constellations with particular conjunctions are more or less likely or not possible (cf. [8], [9]). In the present contribution we present freshly sourced quantitative Slovak SNK-corpus data. We consider the sixteen logically possible tense-aspect constellations, and the Slovak connective items: *odkedy; odvtedy, čo / ako; od chvíle, keď / čo / ako; od tých čias, čo / ako; od tej doby, čo / ako*. This quantitative data study is intended to pinpoint the areas of future research; for this purpose at certain instances comparisons are made with Polish, the only other language we have such data for to date.

Keywords: conjunction, tense, aspect, anteriority, simultaneity, taxis, Slovak, Polish

1 INTRODUCTION

This contribution presents the finding of our investigation into the Slovak correspondences of Dutch and English temporal conjunctions, respectively *sinds/sedert* and *since* and the tense-aspect (hereafter TA) constellations of the complex sentences they appear in. The underlying research is part of the ongoing taxis project of the research group “Comparative Slavic Verbal Aspect” at the university of Amsterdam.¹ Our research group usually work within a cognitive-structuralist framework. Earlier research on this particular conjunction included Czech, Polish and Russian.

¹ Cf. <https://aclc.uva.nl/content/research-groups/comparative-slavic-verbal-aspect-and-related-issues/comparative-slavic-verbal-aspect-and-related-issues.html?origin=8ZtCo6MjS%2B6atiMzaszh6A>.

1.1 Setting the scene

The following example is in a few ways very typical of the kind of complex sentence we are dealing with.

- (1) *Odkedy ťa stretol, všetko sa zmenilo.* [LŠti3]²
Since you.ACC meet.M.3SG.PST.PFV, everything REFL change.N.3SG.PST.PFV³
'Since he (had/has) met you, everything (had/has) changed.'⁴

We see here the three noteworthy distinctive elements.

- the Secondary Clause (hereafter SC) *Odkedy ťa stretol* with conjugated verb *stretol* and
- 'since'-connective item *odkedy*;
- the Main Clause (hereafter MC) with conjugated verb *zmenilo*.

Each of the two conjugated verbs introduce TA-meaning, which interacts between SC and MC, but also probably the selected connective item.

A basic semantic analysis of this type of construction based on the earlier research into Dutch, English, Russian and especially Polish (all earlier references) is provided here for a better understanding as it probably will largely coincide in its generalities, although such an analysis is not the main focus of the current contribution, and a lot more can probably be said about that once the Slovak samples have been thoroughly scrutinized for that purpose. The invariant that has been established, then, consists of the following elements:

- The *sinds/since*-connective item introduces an SC-event which starts in the past;
- The connective item carries a sense of anteriority, which has to do with the beginning in the past;
- The SC-event sets, one might say, "opens" a temporal frame (the "SC-frame"), which stretches from that beginning in the past up to and including the deictic center, which may be at the moment of utterance or before, but which need not be "filled" entirely/throughout with the SC-event itself;
- The MC-event takes place in the temporal frame set by the SC (although it need not "cover" it exactly);
- There is also a sense of simultaneity in this construction, which comes about as the MC-event takes place against the SC-temporal frame (= at some time in that temporal frame).

² The source references are verbatim as they are provided by the Slovak National Corpus – prim-8.0-public-sane (hereafter SNK).

³ Our interlinear glosses follow the Leipzig Glossing Rules. (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>). See also our list of abbreviations below.

⁴ Please note that in the glosses and translation of examples we have not wanted to pinpoint the exact English aspect-forms. That would be quite impossible and is very dependent on context and a few other factors, which have no bearing on the Slovak originals.

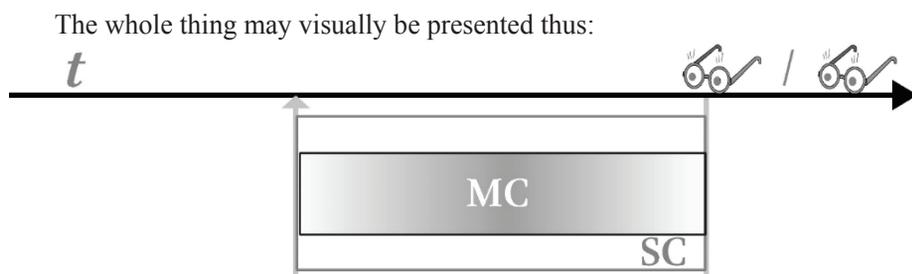


Fig. 1. Invariant meaning of *sinds/since* constructions (after [9] for Polish)

The spectacles represent the location from which an event is observed and it may or may not coincide with the moment of utterance or come after, depending on the particular TA-constellation used.

1.2 Sample examples

The following is a list of examples with more or less randomly chosen *sinds/since*-connective items – it was not thought necessary in this paper with its main focus on the quantitative data to provide an example of each TA-constellation for each connective item, although the exchangeability of the items has not yet been studied. You will notice that passive forms have been deselected here, and choices are intended to be as “riskless” as possible, again, because the deeper semantic analysis is beyond the scope of this paper and the following provides a sufficient impression for our present purposes.

Please note, that in Dutch and English (for which latter, cf. [10, p. 91] the temporal location of the SC-event may not be placed in its entirety in the future. SC-future tenses have not been encountered for Polish and for Slovak they have been omitted here; moreover, they are not likely to fall within the scope of *sinds/since* and one of our purposes is a comparison with Polish. Please note in this respect that PRS. PFV is not qualified here as future, although it deictically often functions as such.

	SC		MC		example (ex SNK)
1	PST	PFV	PST	PFV	(1) <i>Odkedy ťa stretol, všetko sa zmenilo.</i> [LŠti3] 'Since he (had/has) met you, everything (had/has) changed.' ⁵
2	PST	PFV	PST	IPFV	(2) <i>Od chvíle, keď odišla sanitka, pri Holmbergových nohách sa chúlil Wurst.</i> [KBlo1] 'From the moment the ambulance left, Wurst snuggled to Holmberg's feet.'

⁵ Please note that the English translations in this table are for working purposes only: their TA-constellations will depend on matters at play in English that are beyond the scope of our research here. Only in example (1) have we given a few alternatives to give an impression. Please, note also that example (1) is the same as provided above and so we kept the same number.

	SC		MC		example (ex SNK)
3	PST	PFV	PRS	IPFV	(3) <i>Sledujeme vás od chvíle, čo ste vkročili do močiara</i> [JČerv3] 'We've been following you since you stepped into the swamp.'
4	PRS	IPFV	PRS	IPFV	(4) <i>Od tej doby, ako hrávam futbal, viem čo mám robiť.</i> [MYTR2009/01] 'Since [*I play] I have been playing football, I know what to do.'
5	PST	IPFV	PST	PFV	(5) <i>S nostalgiou si uvedomil, že odvtedy, čo bol bezstarostný chalan, zmenilo sa mnoho vecí.</i> [EMcb11] 'With nostalgia, he realized that many things had changed since he was a carefree boy.'
6	PST	IPFV	PST	IPFV	(6) <i>Prvý raz odvtedy, čo žil na veľchánovom dvore, nesprevádzal cisára na letné sídlo.</i> [WMei1] 'For the first time since he had been living at court, he did not escort the emperor to the summer residence.'
7	PRS	IPFV	PST	PFV	(7) <i>Od tej doby, čo sa zúčastňujem tohto projektu, som sa naučila vážiť si samu seba.</i> [ASP2002/01] 'Since I attended this project, I have learned to appreciate myself.'
8	PST	PFV	PRS	PFV	(8) <i>A odvtedy, čo ušli trpaslíci, nikto sa neodváži prehľadávať šachty a poklady v hĺbinách.</i> [JTol2] 'And since the dwarves have gone, no one dares to search the shafts and treasures in the depths.'
9	PRS	IPFV	PST	IPFV	(9) <i>Nielen bohovia, ale aj ľudia vždy potrebovali smiech a zadovažovali si ho od tých čias, čo tu existujú, lenže veľmi primitívne...</i> [ABed8] '[...] but also the people always needed laughter and they obtained it since they existed here [...] ...'
10	PST	IPFV	PRS	IPFV	(10) <i>Ved' odvtedy, ako účinkovali v našom programe, nepretržite pracujú až dodnes.</i> [MYNO2016/09] 'After all, ever since they participated in our program, they have been working [in work] uninterruptedly until today.'
11	PRS	IPFV	PRS	PFV	(11) <i>Odvtedy, čo navštevujem Grécko, vždy si priveziem 5 litrov olivového oleja, ...</i> [MYŽN2009/31] 'Since I have been visiting Greece, I every time bring 5 liters of olive oil, ...'
12	PST	IPFV	PRS	PFV	(12) <i>... ešte vlani ho piekli každý druhý deň, ale odvtedy, čo šiel starý Nérer do penzie, zoženieš ho už len v piatok)...</i> [JJoh2] '... Last year they bake it every other day, but since old Nerer went to the pension, you will only get it on Friday) ...'

Tab. 1. Sample examples of all researched TA-constellations – random connective items

These 12 examples cover all the basic TA-constellation types and the following further examples are provided to cover the special types that are mentioned for

Polish [8] and have now been recognized in Slovak. To complete this issue, we will briefly discuss these.

In both the Polish and now also the Slovak dataset, many examples of various TA-constellations occurred with the meaning ‘remember’.

- (13) *Ja od chvíle, čo si pamätám, som vždy cítil potrebu bojovať.* [HN2011/05]
‘Me, ever since I can remember / for as long as I can remember, I have always felt the need to fight.’

These are translatable by a *sinds/since* construction but there the alternative ‘for as long as’ is usually lacking for other verbs. In terms of taxis this is a noteworthy difference.

Another frequent subtype, for which its distribution across the TA-constellations in Slovak still needs to be studied, might be described as ‘the passing of time’, cf.(14).

- (14)a *Zisťuje, že ubehlo asi dvadsať minút od chvíle, keď zdriemol.* [LT1998/07]
‘He discovers that about twenty minutes have passed since he took a nap.’

A perhaps somewhat unexpected type has future relevance, but is nevertheless translatable with *since*.

- (14)b *Dvadsiateho siedmeho decembra uplynie mesiac odvtedy, čo má sadru.* [KOR2001/12]
‘On the twenty-seventh of December, a month will have passed since he has a cast.’

Not at all surprising is the fact that phase verb ‘begin’ crops up often in this dataset; the connective’s meaning ‘since’ marks a point at which an event commences. Example (15) is rather special as it has such a phase verb both in SC as in MC.

- (15) *Narodí sa nám dieťa, no nie dieťa, ktoré jednoducho začne umierať od chvíle, keď začína žiť.* [KRah1]
‘A child was born unto us, but not a child who simply begins to die from the moment he begins to live.’

The last of our special mentions is an issue touched upon already by [10] and [7] for English and by o.a. [8], [11] and [12] for Slavic languages. It is the issue of the necessity in many examples for an element in the MC that gives some sort of “weight”, some specific “load” or relevance, without which a sentence may be grammatically well-formed, but is nevertheless bad. In (16) the relevant element is underscored.

- (16) *Myšlienka na odchod z ostrova mi neskrsla v hlave ani raz od chvíle, čo som na ňom pristál.* [YMar1]
'The thought of leaving the island didn't occur to me even once, since I came upon it.'

This concludes the samples and we shall leave the remainder of the semantic analysis for future research: the reader at this point should have a sufficient impression of what is involved with the *sinds/since* temporal constructions.

In this paper we proceed by describing the methods used for identifying Slovak correspondences to Dutch *sinds* and English temporal *since* using the *ASPAC* parallel aligned corpus and for compiling our data set using the SNK. Further on we present the quantitative data: on that basis we will discuss the statistical analysis and propose a few lines of research to follow up, as ultimately the (future) goal is to make an extensive semantic analysis and establish the set of usage restrictions for each of the individual connective items.

2 DATA SET, METHODOLOGY

2.1 Identification of the correspondences

As mentioned in 1. the starting point for this research were the Dutch connective item *sinds* and English *since* in temporal usage.⁶ As for our earlier research on Polish, *ASPAC* – Amsterdam Slavic Parallel Aligned Corpus⁷ was used: simple queries via ParaConc yielded the correspondences: *odkedy; odvtedy, čo; odvtedy, ako; od chvíle, ked'; od chvíle, čo; od chvíle, ako; od tých čias, čo; od tých čias, ako; od tej doby, čo; od tej doby, ako.*

Hereafter we will refer to these as 'connective items' rather than as 'conjunctions' on account of their sometimes complex shape, consisting of more than one lexical element, which are sometimes exchangeable. In this respect it is important to note that this method to identify the connective items deselects such items that fall outside the (semantic and other) scope of the connective items in the source languages. This is important as it has been established ([8], [9]) that at least the Polish connective items are not primarily restricted to use with past deixis such as their Dutch and English counterparts. It was expected – and has indeed been established, cf. the following – that this holds for the Slovak counterparts as well.

⁶ The English connective items under scrutiny are *since* and *ever since* in their temporal functions only. As was pointed out in [6], [7], [8], [9] and [10] non-temporal *since* such as we find in e.g. *Since I'm a taxi driver, I know how to get there* allows for TA-constellations that are at variance with those for temporal (*ever*) *since*. Dutch *sinds* is always temporal.

⁷ *ASPAC* is a non-tagged, restricted access corpus compiled by Adrie Barentsen. It consists of original and translated Slavic literary texts. Cf. <http://www.uva.nl/profiel/b/a/a.a.barentsen/a.a.barentsen.html>. We are indebted to Barentsen for the access provided.

2.2 Strategies for compiling the data set with the SNK

Our data, once sourced from the SNK, was set out onto a simple Excel-sheet. It was our purpose to classify all corpus search results per connective item and TA-constellation of MC and SC taken together. Of the sixteen logically possible combinations, twelve were actually encountered and found to equate *sinds/since* usage and they are plotted onto tables 1 through 3. The four possible constellations with SC:PRS.PFV only yielded a very limited number of samples that (for various reasons) never correspond to *sinds/since*. Obviously the expected future deixis of SC:PRS.PFV, often sets a time frame in the future, which is incompatible with *sinds/since*. There are other, not yet researched instances as well: cf. (17) in which there is a kind of generic/repetitive (exemplary) meaning intended.

(17) *Väčšina ľudí si o vás vytvorí prvý dojem najneskôr do piatich sekúnd od chvíle, ako vás zbadá.* [InZ2000/02]

Most people DAT about you make.3SG.PRS.PFV first impression at_most to five seconds from moment that you.ACC notice.3SG.PRS.PFV.

‘Most people will have made a first impression within five seconds from the moment they will (have) notice(d) you.’

As such these types have been deselected from this research.

As we were searching the SNK for our connective items, it soon became clear that some search strings yielded many thousands of hits (e.g. 16067 for *odkedy* alone), whilst others amounted to just a tiny handful of examples. It was decided early on that it would be more practical to have a different treatment for the highly frequent connective items as opposed to the rather lower scoring items; the material for the lower scoring items – *od chvíle, ako; od tých čias, čo; od tých čias, ako; od tej doby, čo; od tej doby, ako* – counted “manually”, which means that the TA constellations for both MC and SC were classified by the researchers for all hits without searches for tagged grammatical categories. We considered this to be the most accurate way to deal with material with low scores.

The higher scoring connective items – *odkedy; odvtedy, čo; odvtedy, ako; od chvíle, ked; od chvíle, čo* – needed to be sourced, classified and counted via SNK-search strings with specified TA for both MC as well as SC. For this purpose we used the fact that the SNK corpus is tagged for tense as well as for aspect: verbs in PST and PRS are tagged respectively VL.* and VK.*, and PFV and IPFV are tagged V.d.* and V.e.* respectively. The search was done in two tiers to cover the two possible sequence-types of the clauses: SC-MC and MC-SC. The queries for *odvtedy, čo*, then, are as follows:

Type 1: clause sequence SC:PST.PFV-MC:PST.PFV:

Odvtedy, čo odišli zo Slovenska, už ubehlo desať rokov. [SME2014/07]

‘Since they (had/have) left Slovakia, ten years (had/have) passed.’

- Search CQL: [lemma="odvtedy"] [lemma=","] [lemma="čo"] within <s/>

- Filter: positive, range (0,10) incl KWIC, CQL: [tag="VLd.*"] [] {0,5} [word="",] [] {0,5} [tag="VLd.*"]⁸

Type 2: clause sequence MC:PST.PFV-SC:PST.PFV:

Prešiel už rok odvtedy, čo ocko oslávil päťdesiatiny. [EFBS1]

‘A year (has/had) already passed since father (has/had) turned fifty.’

- Search CLQ: [lemma="odvtedy"] [lemma="",] [lemma="čo"] within [tag="VLd.*"] [] {0,10} [tag="VLd.*"] within </s>
- Filter: negative, range (-1,-1) excl. KWIC, CQL: lemma = ""

(The filter is applied to ensure the deselection of sentences that begin with the queried string itself when this belongs to the previous sentence in the text.)

For each of the logically possible TA-constellations as well as for each of the identified connective items the above search strings were adjusted.

In the further processing, the “order-types” were not separated out in any of the calculations as the clause order was deemed inconsequential for the present research (although we do not exclude it could be worth researching these variations at a later stage). The totals – and indeed all working figures of the higher scoring items – provided in table 1 are then the sum of the scores for these two order-types. The SNK search options allow for other solutions to our problems, but this turned out to be easy enough to apply successfully, although we realize that an “automatic” count is never as precise as the “manual” one, we applied for the lower scoring connective items. Some existing false positive or negative results do affect accuracy of the results but for totals in the order of thousands we consider the statistical error to be low enough for this early study. Further research is needed to establish the exact inaccuracy of similar corpus queries.

In the following we will present the quantitative data.

nr	SC		MC		sum of all connective items		sum of all connective items – odkedy		odkedy	
	T	A	T	A						
1	PST	PFV	PST	PFV	6456	26.04%	2884	32.90%	3572	22.28%
2	PST	PFV	PST	IPFV	5493	22.15%	2075	23.67%	3418	21.32%
3	PST	PFV	PRS	IPFV	4119	16.61%	1382	15.76%	2737	17.07%
4	PRS	IPFV	PRS	IPFV	1859	7.50%	240	2.74%	1619	10.10%

⁸ The limitations on the word range (= 10) were set after some experimenting with other ranges. Shorter ranges excluded many usable examples whilst wider ranges introduced inaccuracies in the scores on account of interference from other than the targeted verbs in the sentences.

nr	SC		MC		sum of all connective items		sum of all connective items – odkedy		odkedy	
	T	A	T	A						
5	PST	IPFV	PST	PFV	1615	6.51%	751	8.57%	864	5.39%
6	PST	IPFV	PST	IPFV	1431	5.77%	419	4.78%	1012	6.31%
5	PRS	IPFV	PST	PFV	1240	5.00%	87	0.99%	1153	7.19%
7	PST	PFV	PRS	PFV	888	3.58%	634	7.23%	254	1.58%
8	PRS	IPFV	PST	IPFV	834	3.36%	33	0.38%	801	5.00%
9	PST	IPFV	PRS	IPFV	659	2.66%	234	2.67%	425	2.65%
11	PRS	IPFV	PRS	PFV	168	0.68%	10	0.11%	158	0.99%
12	PST	IPFV	PRS	PFV	35	0.14%	18	0.21%	17	0.11%
totals:					24797	100%	8767	100%	16030	100%

nr	SC		MC		odvtedy, čo		odvtedy, ako		od chvíle, keď	
	T	A	T	A						
1	PST	PFV	PST	PFV	1407	32.32%	499	29.72%	455	40.23%
2	PST	PFV	PST	IPFV	716	16.45%	381	22.69%	400	35.37%
3	PST	PFV	PRS	IPFV	709	16.29%	352	20.96%	102	9.02%
4	PRS	IPFV	PRS	IPFV	113	2.60%	67	3.99%	9	0.80%
5	PST	IPFV	PST	PFV	460	10.57%	115	6.85%	103	9.11%
6	PST	IPFV	PST	IPFV	241	5.54%	84	5.00%	21	1.86%
7	PRS	IPFV	PST	PFV	34	0.78%	38	2.26%		
8	PST	PFV	PRS	PFV	507	11.65%	72	4.29%	22	1.95%
9	PRS	IPFV	PST	IPFV	18	0.41%	5	0.30%	1	0.09%
10	PST	IPFV	PRS	IPFV	138	3.17%	61	3.63%	7	0.62%
11	PRS	IPFV	PRS	PFV	6	0.14%		0.00%	3	0.27%
12	PST	IPFV	PRS	PFV	4	0.09%	5	0.30%	8	0.71%
totals:					4353	100%	1679	100%	1131	100%

nr	SC		MC		od chvíle, čo		od chvíle, ako		od tých čias, čo	
	T	A	T	A						
1	PST	PFV	PST	PFV	218	44.86%	159	27.46%	82	24.19%
2	PST	PFV	PST	IPFV	175	36.01%	264	45.60%	88	25.96%
3	PST	PFV	PRS	IPFV	23	4.73%	109	18.83%	44	12.98%
4	PRS	IPFV	PRS	IPFV	6	1.23%	7	1.21%	24	7.08%

nr	SC		MC		<i>od chvíle, čo</i>		<i>od chvíle, ako</i>		<i>od tých čias, čo</i>	
	T	A	T	A						
5	PST	IPFV	PST	PFV	21	4.32%	9	1.55%	32	9.44%
6	PST	IPFV	PST	IPFV	15	3.09%	18	3.11%	37	10.91%
7	PRS	IPFV	PST	PFV			2	0.35%	10	2.95%
8	PST	PFV	PRS	PFV	24	4.94%	8	1.38%		
9	PRS	IPFV	PST	IPFV	1	0.21%		0.00%	5	1.47%
10	PST	IPFV	PRS	IPFV	3	0.62%	3	0.52%	15	4.42%
11	PRS	IPFV	PRS	PFV				0.00%	1	0.29%
12	PST	IPFV	PRS	PFV				0.00%	1	0.29%
totals:					486	100%	579	100%	339	100%

nr	SC		MC		<i>od tých čias, ako</i>		<i>od tej doby, čo</i>		<i>od tej doby, ako</i>	
	T	A	T	A						
1	PST	PFV	PST	PFV	36	30.77%	22	30.99%	6	50.00%
2	PST	PFV	PST	IPFV	32	27.35%	19	26.76%		
3	PST	PFV	PRS	IPFV	24	20.51%	15	21.13%	4	33.33%
4	PRS	IPFV	PRS	IPFV	5	4.27%	7	9.86%	2	16.67%
5	PST	IPFV	PST	PFV	10	8.55%	1	1.41%		
6	PST	IPFV	PST	IPFV	3	2.56%		0.00%		
7	PRS	IPFV	PST	PFV	2	1.71%	1	1.41%		
8	PST	PFV	PRS	PFV	1	0.85%		0.00%		
9	PRS	IPFV	PST	IPFV	3	2.56%		0.00%		
10	PST	IPFV	PRS	IPFV	1	0.85%	6	8.45%		
11	PRS	IPFV	PRS	PFV				0.00%		
12	PST	IPFV	PRS	PFV				0.00%		
totals:					117	100%	71	100%	12	100%

Tab. 2. Quantitative data sourced from the SNK, processed in Excel

3 DISCUSSION OF THE QUANTITATIVE DATA

3.1 Slovak connective item

The numbers of SNK-hits were set out onto an Excel sheet and calculations were made to produce totals and percentages. Table 2 shows the results of this rounded off to two decimal places. The connective items are set out according to

frequency from left to right. The TA-constellations across SC and MC are sorted according to frequency from top to bottom of the total occurrence of each TA-constellation. In this way the preponderance/preference of each connective item per TA-constellation can be easily judged. For better legibility we have refrained from marking zero (and 0.00%) scores.

It will be apparent that of the logically possible TA-constellations four are missing and this was discussed in 2.2 above.

The data shows that *odkedy* is by far the most highly scoring connective item not only in total, but also in all constellation types, and even, albeit with the notable exception of type 7 (and less notably 12), in absolute numbers against the sums of the scores of the other connective items taken together. Its frequency of occurrence, however, does not trace that of the totals for all connective items together – even though those totals are of course heavily influenced by *odkedy* itself. For this reason as well as its unevenly high score we have produced a further column with totals for all connective items except *odkedy*. The differences between *odkedy*-scores against this column are even greater – as expected – and they display some interesting features which will be pointers for future research. For now we want to point out that our research has so far not produced any evidence that the respective connective items should not be exchangeable in any clear way and so we cannot at this point in time pinpoint why, in what circumstances a particular connective item is preferred. Here we can only mention in which TA-constellation there are significant deviations of *odkedy* against the others. The following table 3 shows this.

nr	SC		MC		sum of all connective items – <i>odkedy</i>	<i>odkedy</i>
	T	A	T	A		

others outscored by *odkedy*

4	PRS	IPFV	PRS	IPFV	240	2.74%	1619	10.10%
7	PRS	IPFV	PST	PFV	87	0.99%	1153	7.19%
9	PRS	IPFV	PST	IPFV	33	0.38%	801	5.00%

odkedy outscored by others

8	PST	PFV	PRS	PFV	634	7.23%	254	1.58%
---	-----	-----	-----	-----	-----	-------	-----	-------

Tab. 3. Quantitative data sourced from the SNK, processed in Excel

Apart from the anyway very low scoring *od tej doby, čo* and *od tej doby, ako*, the other items have a preponderance for especially type 5 (SC:PST.IPFV-MC:PST.PFV) and most, but notably not *od tých čias, čo*; *od tých čias, ako*; *od tej doby, čo*; *od tej doby, ako* also for type 8 (SC:PST.PFV-MC:PRS.PFV).

The choice of connective item seems to only be very slightly influenced or determined by the TA-constellation with only a few “preferences”. To date we have, pending our semantic analysis of the data, not conclusively established why these preferences occur and why the distribution of all connective items is not more even, especially that of *odkedy*.

3.2 Sum of all connective items: Slovak & Polish

We took the sum of the occurrences in our data set of each of the researched TA-constellations with SC+MC taken together and compared that data with similar data for Polish, the only language for which such data is at hand. As the purpose of this research was to establish line of further research, it seemed relevant to do so and learn what we could: significant differences would pinpoint areas for further research. That yielded table 4.⁹

nr	SC		MC		Slovak		Polish	
	T	A	T	A				
1	PST	PFV	PST	PFV	6456	26.04%	431	31.23%
2	PST	PFV	PST	IPFV	5493	22.15%	324	23.48%
3	PST	PFV	PRS	IPFV	4119	16.61%	241	17.46%
4	PRS	IPFV	PRS	IPFV	1859	7.50%	115	8.33%
5	PRS	IPFV	PST	PFV	1240	5.00%	52	3.77%
6	PST	IPFV	PST	IPFV	1431	5.77%	61	4.42%
7	PST	IPFV	PST	PFV	1615	6.51%	70	5.07%
8	PRS	IPFV	PST	IPFV	834	3.36%	51	3.70%
9	PST	IPFV	PRS	IPFV	659	2.66%	21	1.52%
10	PST	PFV	PRS	PFV	888	3.58%	12	0.87%
11	PRS	IPFV	PRS	PFV	168	0.68%	2	0.14%
12	PST	IPFV	PRS	PFV	35	0.14%	0	0.00%
totals:					24797	100%	1380	100%

Tab. 4. Sum of all TA-constellations Slovak and Polish [9]

It is rather satisfying – although not entirely unexpected – that Slovak and Polish display virtually the same percentages per TA-constellation. The deviations have been marked in the table by the outline and this concerns types 5, 6 and 7 but they are very slight indeed.

⁹ This data is sourced from [9], who also presented scores for SC:PRS.PFV items. The latter has been omitted in this table, which presents, then, a direct comparison of the TA-constellation types for the complex sentences for these two languages.

Of course, the Polish dataset was considerably smaller, but that should not deter us from suggesting that these two languages have a very similar treatment in terms of TA-constellation of the *sinds/since* construction.

3.3 Comparing SC and MC scores: Slovak & Polish

Next, we considered the TA-data per clause and so with the scores and percentages for all connective items taken together. In the following two tables 5 the scores for all SC have been summed regardless of the MC they are combined with. Also the tables display the scores for all MC regardless of their SC.

SC				MC			
T	A	score	perc.	T	A	score	perc.
PST	PFV	16956	68.38%				
PRS	IPFV	4101	16.54%				
PST	IPFV	3740	15.08%				
<i>PRS</i>	<i>PFV</i>	<i>omitted</i>	-				
				PST	PFV	9311	37.55%
				PST	IPFV	7758	31.29%
				PRS	IPFV	6637	26.77%
				PRS	PFV	1091	4.40%
control total:		24797	100%			24797	100%

Tab. 5a. SC resp. MC: Slovak

This particular Slovak data can at this time be compared to similar quantitative data for Polish, the only other language for which it is available even though the latter is based on considerably smaller quantities of samples. The figures provided by [9] have been adapted to the following table formatted for easy comparison.

SC				MC			
T	A	score	perc.	T	A	score	perc.
PST	PFV	1008	72.41%				
PRS	IPFV	220	15.80%				
PST	IPFV	152	10.92%				
PRS	PFV	12	0.86%				
				PST	PFV	553	39.73%
				PST	IPFV	438	31.47%
				PRS	IPFV	387	27.80%
				PRS	PFV	14	1.01%
control total:		1392	100%			1392	100%

Tab. 5b. SC resp. MC: Polish (cf. [9])

In spite of the difference in sample volumes as well as the fact that for Slovak SC:PRS.PFV has been omitted,¹⁰ we may at this time, however, signal that the scores of the types in percentages are very similar indeed.

4 CONCLUSIONS

As stated above in the introduction, the purpose of this paper is to set out further lines of research concerning Slovak correspondences of Dutch/English temporal *sinds/since* sentences: TA-constellations and connective items. We have utilized the *ASPAC* and the *SNK* corpora to compile a dataset and on the quantitative data thus at hand, we propose the following notes in respect to future research:

- Both in Polish and Slovak the types with PRS.PFV (be it in SC or MC) are rare. Although in these languages the *sinds/since* correspondences were found not to be restricted to past deixis, nevertheless it would seem that connective items such as those under discussion still have a prepondering use for past deixis. The PRS.PFV types need to be further researched to establish the extent of their usage as well as the functioning of the TA-constellations. NB. The particular “exemplary” use of PRS.PFV in (17) certainly needs further research and future deixis does not seem obvious.

This will aid formulating the very definition of the Slovak (and indeed Polish) connective items and their usage.

- The comparison of Slovak vs. Polish data shows that these two languages do not differ significantly in TA-constellations encountered for *sinds/since* complex sentences. There are a few points at which there are small differences and one might surmise that these are due to as yet not researched differences in the respective TA systems of these languages.
- The meaning types of the TA-constellations need to be established more precisely for Slovak (such as has been done e.g. for Polish in [9]).
- The preference(s) for particular TA-constellations by the individual connective items needs further research. An inroad into this will be the outlying scores in our table 2 and also 3. This will further establish the preferred meaning type(s) but also pinpoint the meaning and usage of the connective items.
- In our research to date on Slovak we have not considered register and clause order (SC-MC vs MC-SC) as factors in the selection of particular connective items. There are, however, some indications from closely related Czech set out by [5] that such matters are not without their significance in the choice of connective item. This still needs to be followed up for Slovak.

In future contributions the present authors intend to address some of these matters, especially the semantic analysis.

¹⁰ Although precise counts are not available at this time, it is clear that they are very slight indeed.

ABBREVIATIONS

3SG = third person singular; ACC = accusative; ASPAC see below under corpora; DAT = dative; IPFV = imperfective; M = masculine; MC = main clause; N = neuter; PFV = perfective; PST = past tense; REFL = reflexive pronoun; SC = secondary clause; SNK see below under corpora; TA = tense-aspect.

References

- [1] Barentsen, A. (2009). Taxis v niderlandskom jazyke. In *Tipologija taksisnyx konstrukcij*, pages 269–366, V. S. Xrakovskij. Moskva, Znak.
- [2] Barentsen, A. (2016). Taxis in Dutch. In *Typology of Taxis Constructions (= LINCOM Studies in Theoretical Linguistics 58)*, pages 205–276, V. S. Xrakovskij. München, Lincom.
- [3] Barentsen, A. (2017). Nekotorye nabljudenija o niderlandskom i anglijskom sojuzax sinds/ since i ix sootvetstvijax v sovremennyx slavjanskix jazykax. In *Definitely Perfect: Festschrift for Janneke Kalsbeek (= Pegasus Oost-Europese Studies 29)*, pages 21–56. Eds. R. M. Genis, E. de Haard and R. Lučić. Amsterdam, Uitgeverij Pegasus.
- [4] Barentsen, A. (2018). O funkcionirovanii vidovremennyx form v složnyx predloženijax s sojuzom tipa s tex por kak / depuis que: na materiale slavjanskix i neslavjanskix jazykov. In *La relation temps/aspect: approches typologique et contrastive (Collection travaux et recherches UL3 / Éditions du Conseil Scientifique de l'Université de Lille 3)*, pages 153–157, T. Milliaressi. Lille, Université Charles-de-Gaulle – Lille3.
- [5] Duijkeren-Hrabová, M. van (2017). Het probleem van variatie bij Tsjechische voegwoorden met de betekenis ‘sinds’. In *Definitely Perfect: Festschrift for Janneke Kalsbeek (= Pegasus Oost-Europese Studies 29)*, pages 107–128. Eds. R. M. Genis, E. de Haard and R. Lučić. Amsterdam, Uitgeverij Pegasus.
- [6] Fijn van Draat, P. (1903). The loss of the prefix ge- in the modern English verb and some of its consequences – The conjunction since. *Englische Studien*, 32, pages 371–388.
- [7] Fryd, M. (2011). Since when is the Present Tense ruled out with SINCE! *Groninger Arbeiten zur germanistischen Linguistik*, 53(2), pages 89–103.
- [8] <https://ugp.rug.nl/GAGL/article/view/30529/27829>.
- [9] Genis, R. M. (2018). Towards a semantic model for conjunctions of ‘since’ and its interaction with verbal aspect: Polish material. In *La relation temps/aspect: approches typologique et contrastive (Collection travaux et recherches UL3 / Éditions du Conseil Scientifique de l'Université de Lille 3)*, pages 159–163, T. Milliaressi. Lille, Université Charles-de-Gaulle – Lille3.
- [10] Genis, R. M. (2019 to appear). Temporal ‘since’ in Polish: conjunctions & tense-aspect constellations. In *Dutch Contributions to the Sixteenth International Congress of Slavists – Linguistics: Belgrade, Belgrade, August 22–27, 2018 (= Studies in Slavic and General Linguistics 44)*, pages [to appear]. Leiden, Brill – Rodopi.
- [11] Heinämäki, O. (1978). *Semantics of English Temporal Connectives*. Indiana, Indiana University Linguistics Club.
- [12] Popović, Lj. (2012). Kontrastivna gramatika srpskog i ukrajinskog jezika: taksis i evidencijalnost. Beograd, Srpska akad. nauka i umetnosti.

- [13] Zorixina-Nil'sson, N. V. (2011). Semantika perfektnosti i dlitel'nosti i sposoby ee vyraženiya v vyskazyvaniyax s sojuzom s tex por kak (Iz sopostavitel'no-tipologičeskix nabljudenij). In *Sistemnye svjazi v grammatike i tekste. Materialy čtenij pamjati Ju. A. Pupykina*. 30 aprēlja 2010 g. M. D. Voejkova, A. Ju. Pupykina, M. Ju. Pupykina, pages 81–125. Sankt-Peterburg, Nestor – Istorija.
- [14] ASPAC = Amsterdam Slavic Parallel Aligned Corpus. Information on the corpus accessible at: <http://www.uva.nl/profiel/b/a/a.a.barentsen/a.a.barentsen.html>
- [15] SNK = Slovenský národný korpus – prim-8.0-public-sane. Bratislava, Jazykovedný ústav Ľ. Štúra SAV 2018. Accessible at: <http://korpus.juls.savba.sk>