Radek Čech (University of Ostrava)
radek.cech@osu.cz

Testing of the Transitivity Hypothesis: Double Object Verbs and Aspect in Czech*

1. Introduction

The *Transitivity Hypothesis* (Hopper & Thompson 1980) predicts universal properties of transitive construction. According to Hopper & Thompson, the transitivity rules the behavior of different grammatical categories (e.g. aspect, negation, mode), with the intent that these seemingly disparate properties closely correlate (for more details see section 4).

The aim of the article is to test one particular prediction of the Transitivity Hypothesis (TH) concerning the relationship between double object constructions and aspect in Czech. In particular, the TH predicts a positive correlation between the perfectivity of verbs and their overt double object syntax. It turns out that the predicted correlation is not statistically significant. A new working hypothesis that accounts for the actual data is proposed. As for the methodology, a corpus-based approach (Bybee & Hopper 2001) is adopted, which means that only language in actual use in authentic discourse contexts is used as material for the linguistic study. Therefore, the first step of this study consists of parsing the language material stored in a language corpus: the method proposed by Mukherjee (2005) was used for the detection of verbs which have strong tendency to occur in double object syntax and for their classification (section 2). Next, verbs detected by this methodology were used for the TH testing (sections 4.2 and 4.3)

2. Methodology and language material

Double object verbs are defined syntactically in the present study.[1] As a double object verb (DOV) is considered any trivalent verb that requires a subject (S), a direct accusative object ($O_{d/acc}$), and an indirect dative or accusative object ($O_{i/dat}$ / $O_{i/acc}$) for a complete syntactic complementation. It is necessary for both objects to be realizable as a noun phrases (NPs); this realization [(S) – DOV – $O{:}NP_{i/dat}$ / $O{:}NP_{i/acc}$ – $O{:}NP_{d/acc}$] is called the basic double object construction. If a verb is attested in the basic double object construction in the actual language use, it is also considered as a double object verb in all other forms of complementation. In total, there are seven parsing schemes for identifying double object verbs in the present paper:

(1) **(S) – DOV – $\mathbf{O{:}NP_{dat}}$ – $\mathbf{O{:}NP_{acc}}$**
(2) **(S) – DOV – $\mathbf{O{:}NP_{acc}}$ – $\mathbf{O{:}NP_{acc}}$**
(3) (S) – DOV – $O{:}NP_{dat}$ – O:CLAUSE
(4) (S) – DOV – $O{:}NP_{acc}$ – O:CLAUSE
(5) (S) – DOV – $O{:}NP_{dat}$ – $O{:}V_{inf}$
(6) (S) – DOV – $O{:}NP_{acc}$ – $O{:}V_{inf}$
(7) (S) – $DOV_{passive}$ – $O{:}NP_{dat}$

* The text was created as part of a grant study for the GAČR (Czech Science Foundation) No. 405/08/P157, *Components of transitivity analysis of Czech sentences (emergent grammar approach)*.

1 This definition is based on Mukherjee's definition of ditransitive verb (Mukherjee, 2005: 80). I prefer to use the term *double object verbs* rather than *ditransitive verbs* in the present study because a *ditransitivity* is commonly defined semantically (Malchukov et al. 2007), contrary to the present analysis which is focused on syntactic properties of verbs.

In accordance with the definition, patterns 1 and 2 above are considered as the basic double object constructions; if the verb is attested in pattern 1 or 2 in the corpus, it is also considered in other patterns (3-7). Owing to the rich verbal morphology in Czech, which expresses among others, person, number, and gender (in the past tense), the presence of the nominal or pronominal subject is not obligatory in Czech. Therefore, both realizations – with and without the overt nominal or pronominal subject – are taken into account. It is important to add that the reflexive passive forms of verbs are not taken into account because of their specific syntax and semantics.

The data used in this study come from the Prague Dependency Treebank 2.0 (further PDT) (Hajič et. al. 2006). The PDT is a Czech corpus which covers a large amount of texts with interlinked morphological, syntactic, and semantic annotation. For the present purposes, the data annotated on the analytical layer were used. This part of the PDT contains 4,269 documents, 68,496 sentences and about 1,5 millions words/tokens. The PDT consists of articles from newspapers and journals. Consequently, all interpretations are determined by this factor and it is necessary to point out that all conclusions presented in this paper are valid only for this text type; compare Biber et al. (1999:8ff). The other limitation of the present analysis is caused by the PDT annotation which ascribes the predicate function only to the head-clause. Because of the use of the PDT annotation for automatic parsing, only head-clause predicates are calculated. Furthermore, it was not possible to use the PDT annotation as a whole for the purpose of this analysis. That is why a few adjustments of the PDT had to be made; for instance, in the case of coordination and apposition annotation. The parsing was done automatically by using the Python programming language.

3. Parsing results

The methodological approach presented in Section 2 has led to the detection of 1391 sentences in which 297 predicate verb lemmas fit the definition of double object verb. For the appropriate TH testing (sections 4.3 and 4.4) the detected verbs were first appropriately classified. In accordance with the usage-based methodology (Bybee & Hopper 2001), the frequency is considered as a main aspect of classification – two frequency parameters are used for expressing typicality of a double object verb (Mukherjee 2005:83): (1) the overall frequency of a verb in the corpus; (2) the frequency with which a verb occurs in a double object syntax. In short, the typical double object verb has to occur frequently on general and frequently in double object syntax in particular. Following (and slightly modifying) Mukherjee's classification, I distinguish three groups of double object verbs:[2]

1. *typical* double object verbs, which are used very frequently as predicates in general (overall frequency > 140 in 1,5 mil. corpus) *and* also frequently in double object syntax (more than 50% of all occurrences of a given verb);
2. *habitual* double object verbs, which are used fairy frequently as predicates in general (overall frequency 200 < 40 in 1,5 mil. corpus), but not in double object syntax (more than

2 Mukherjee's parameters are used only in the case of distinguishing of double object syntax. Mukherjee's parameters for overall frequency cannot be applied because of the PDT annotation – the PDT annotation ascribes the predicate function just to the head-clause which means that the overall number of observed verbs is much smaller, although the size of the PDT is comparable with the corpus which Mukherjee has used. Moreover, frequency just of the predicate verb seems to be better parameter for distinguishing typicality of double object verbs than overall frequency of the verb (regardless of its syntactic function) because schemata for parsing are defined just for predicate verbs. Consequently, I propose different parameters for overall frequency: for peripheral double object verbs fverb ≥ 40, which means that the verb has to occur at least twice for reaching the 5% limit of double object syntax (the limit of periphery). Proportionally to peripheral limit (in Mukherjee's classification sense), for habitual double object verbs I propose fverb ≥ 140.

5% of all occurrences of a given verb have to be in double object constructions);
3. *peripheral* double object verbs, which are used only sporadically as predicates in general (overall frequency < 40 in 1,5 mil. corpus) and/or which are used only rarely in double object syntax (less than 5% of all occurrences of a given verb have to be in double object constructions).

The application of this criteria leads to the detection of one typical double object verb

(1) *sdělit*
'to inform'

and forty-one habitual double object verbs in the PDT, as table 2 and figure 1 below show.

4. The Transitivity Hypothesis

The Transitivity Hypothesis is one of the influential theories of transitivity which allows to be empirically tested (in Popper's (1959) sense). According to Hopper & Thompson (1980:251), "[t]ransitivity is a crucial relationship in language, having a number of universally predictable consequences in grammar". Furthermore, transitivity is viewed as a property of a sentence which comprises ten components (see table 1 below) – each component involves a different facet of the effectiveness or intensity with which the action is transferred from one participant to another. So, transitivity "can be broken into its component parts (…), they allow clauses to be characterized as MORE or LESS Transitive: the more features a clause has in the 'high' column in 1A–J, the more Transitive it is" (p. 253). The most important aspect of the TH lies in the prediction which hypothesizes the relationships between the components: "If two clauses (a) and (b) in a language differ in that (a) is higher in Transitivity according to any features 1A-J, then, if concomitant grammatical or semantic difference appears elsewhere in the clause, that difference will also show (a) to be higher in Transitivity" (p. 255). Component features should co-vary extensively and systematically, so "whenever two values of the transitivity components are necessarily present (...) they will agree in being either both high or both low in value" (1980:254). Of course, the co-variation has to be viewed not in the strict sense, but as a tendency; cf. Thompson & Hopper (2001).

However, there has not been any satisfactory answer in linguistics to the question concerning the origin of transitivity in this sense until now (Naess 2007:16), so a closer investigation of the TH is still necessary. Moreover, the TH has not, to my knowledge, been tested in Czech. Therefore, the next sections are devoted to the testing of the hypothesis.

Table 1: Transitivity parameters

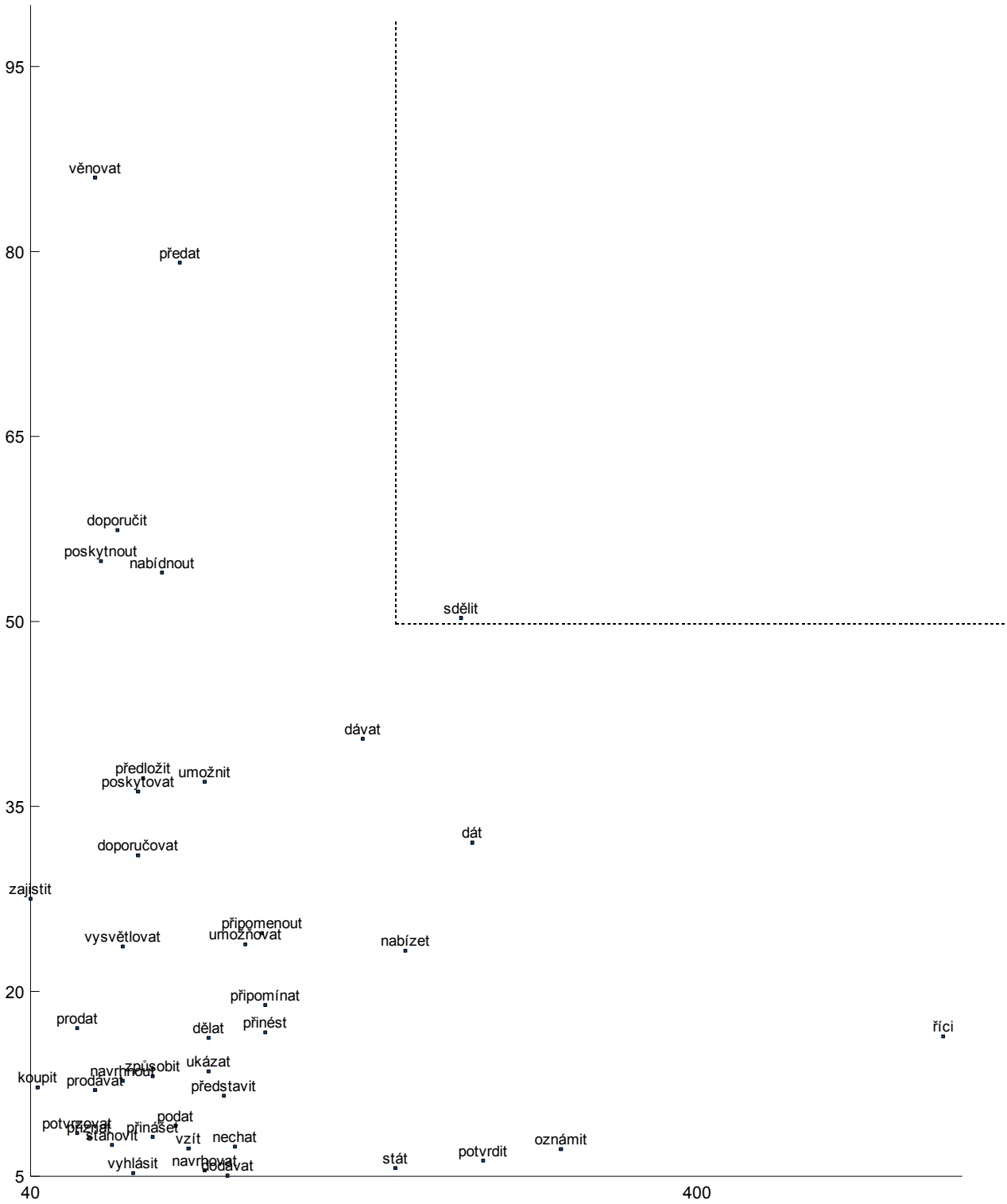
		high T	low T
A	PARTICIPANTS	2 or more	1
B	KINESIS	action	non-action
C	ASPECT	telic	atelic
D	PUNCTUALITY	punctual	non-punctual
E	VOLITIONALITY	volitional	non-volitional
F	AFFIRMATION	affirmative	negative
G	MODE	realis	irrealis
H	AGENCY	A high in potency	A low in potency
I	AFFECTEDNESS of O	O totally affected	O not affected
J	INDIVIDUATION of O	O highly individuated	O non-individuated

Table 2: The list of double object verbs[3]

verb	overall frequency of predicate verb	parsed as double object verb	percentage of double object syntax
věnovat ‘donate’	50	43	86
předat ‘hand over’	67	53	79.1
doporučit ‘recommend’ (perf.)	54	31	57.4
poskytnout ’allow’	51	28	54.9
nabídnout ‘offer’ (perf.)	63	34	54
sdělit ‘inform’	**177**	**89**	**50.3**
dávat ‘give’ (imperf.)	126	51	40.5
předložit ‘submit’	59	22	37.3
umožnit ‘enable’ (perf.)	73	27	37
poskytovat ‘provide’	58	21	36.2
dát ‘give’ (perf.)	184	59	32.1
doporučovat ‘recommend’ (imperf.)	58	18	31
zajistit ‘ensure’	40	11	27.5
připomenout ‘remind’ (perf.)	89	22	24.7
umožňovat ‘enable’ (imperf.)	84	20	23.8
vysvětlovat ‘explain’	55	13	23.6
nabízet ‘offer’ (imperf.)	146	34	23.3
připomínat ‘remind’ (imperf.)	90	17	18.9
prodat ‘sell’ (perf.)	47	8	17
přinést ‘bring’ (perf.)	90	15	16.7
říci ‘tell’	936	153	16.4
dělat ‘do’ (imperf.)	74	12	16.2
ukázat ‘show’	74	10	13.5
způsobit ‘cause’	61	8	13.1
udělat ‘do’ (perf.)	61	8	13.1
navrhnout ‘propose‘ (perf.)	55	7	12.7
koupit ‘buy’	41	5	12.2
prodávat ‘sell’ (imperf.)	50	6	12
představit ‘introduce’	78	9	11.5
podat ‘give’	66	6	9.1
potvrzovat ‘confirm’ (imperf.)	47	4	8.5
přinášet ‘bring’ (imperf.)	61	5	8.2
přiznat ‘award’	49	4	8.2
stanovit ‘assigne’	53	4	7.6
nechat ‘let’	81	6	7.4
vzít ‘take’	69	5	7.3
oznámit ‘announce’	250	18	7.2
potvrdit ‘confirm’ (perf.)	191	12	6.3
stát ‘cost’	141	8	5.7
navrhovat ‘propose’ (imperf.)	73	4	5.5

3 The verbs are ranked according to percentage of the double object syntax. If there are both aspectual forms of a verb (perfective and imperfective) in the list, the aspect of the verb is marked.

Figure 1: The frequency-based distribution of typical and habitual double object verbs[4]



4 The x-axis is logarithmic and expresses the overall frequency. The y-axis expresses percentage of the double object syntax. The dashed line identifies the border line between typical and habitual verbs.

4.1 The Transitivity Hypothesis with regard to the double object syntax

The presence of an object (or objects) is one of the ten parameters which indicate high transitivity of the sentence.[5] The evident differences between properties of direct and indirect object, namely the tendency of an indirect object to be animate and definite, led Hopper & Thompson (1980) to the more precise predictions of the distinct objects behavior. Specifically, both the animacy and definiteness are high transitivity features of the parameter called Individuation of Object (parameter J); so there is a positive correlation between indirect object and this parameter. Therefore, indirect objects "should in fact be Transitive O's [objects] rather than what might be called 'accusative' O's" (Hopper & Thompson 1980:259). In accordance with the approach, Hopper & Thompson generalize and predict "that clauses containing indirect O's will indicate high Transitivity in some other respect" (p. 260). In other words, sentences with the double object construction should be more transitive than sentences with the single object construction.

4.2 The hypothesis testing – the double object constructions and the aspect

The description of the Czech double object verbs, presented in section 3, seems to be a good starting point for the hypothesis testing; the most typical double object verbs (with regard to frequency) are obviously the best candidates for the testing. Moreover, it is possible to observe a prospective relationship between the double object typicality, as defined in section 3, on one hand, and extent of the correlation of parameters, on the other. In brief, we can ask the question: is the correlation between the double object syntax and a chosen high transitive parameter stronger in the case of typical double object verbs than in the case of less-typical double object verbs?

There is not enough space in this paper for testing all parameters, therefore, only the relationship between double object syntax and the aspect is analyzed. The aspect is not chosen accidentally, there are at least two main reasons for choosing an aspect: (1) the aspect is formally well distinguishable in Czech; (2) the aspect is an independent parameter. As for the latter reason, although Hopper and Thompson presented 10 parameters which could correlate only because of transitivity, some of the features are dependent on the others by definition; for details see Olsen & Macfarland (1996). In fact, only the number of participants, kinesis, aspect, affirmation, and mode are independent parameters.

As for the testing itself, the frequency characteristics are observed. Although the frequency approach is not suggested in the original Hopper & Thompson's paper (1980), frequency is used as a main testing parameter for the TH in their later work (Thompson & Hopper 2001). The TH predicts that sentences with a double object construction should correlate with perfectivity. In other words, if we have a pair of verbs which occur in both double object and single object constructions and which only differ in an aspect (parameter C), the hypothesis predicts that perfective verbs have a tendency to occur more often in a double object than single object construction.

The hypothesis was tested as follows:

1. those verbs were chosen for testing which are marked as typical or habitual double object verbs (see section 3), as well as their aspect counterparts,
2. the frequency was observed of the usage of each verb in a double object construction and in a single object construction,

5 But contrary to most approaches, the presence of the object in the sentence is not the necessary condition for consideration of the *transitivity* of the verb – it is possible that a sentence without the object could be "more" transitive than a sentence with the object, because higher transitivity parameters could be present in the sentence without object; see Hopper & Thompson (1980:254).

3. the results of each pair of verbs, both perfective and imperfective, were compared (see table 3) and differences tested by the Fischer's exact test.

4.3 Results

There are 11 pairs of testable typical or habitual double object verbs in the PDT, as shown in table 3.

Table 3: The occurrences of verbs with a double object and single object syntax[6]

verb	aspect	ditransitive	monotransitive	percentage of ditransitive syntax	Fischer's test p-value
předat 'hand over'	perf.	53	12	82	0.153
předávat 'hand over'	imperf.	4	3	57	
doporučit 'recommend'	perf.	31	23	57	**0.012**
doporučovat 'recommend'	imperf.	18	38	32	
nabídnout 'offer'	perf.	34	28	55	**0.00001**
nabízet 'offer'	imperf.	34	112	23	
poskytnout 'allow'	perf.	28	23	55	0.056
poskytovat 'allow'	imperf.	21	37	36	
sdělit 'inform'	perf.	89	88	50	1
sdělovat 'inform'	imperf.	3	4	43	
dát 'give'	perf.	59	119	33	0.339
dávat 'give'	imperf.	51	69	43	
předložit 'submit'	perf.	22	37	37	0.562
předkládat 'submit'	imperf.	7	8	47	
umožnit 'enable'	perf.	27	46	37	0.082
umožňovat 'enable'	imperf.	20	64	24	
připomenout 'remind'	perf.	22	50	31	0.098
připomínat 'remind'	imperf.	17	73	19	
vysvětlit 'explaine'	perf.	9	22	29	0.438
vysvětlovat 'explaine'	imperf.	11	41	21	
zajistit 'ensure'	perf.	11	29	28	0.146
zajišťovat 'ensure'	imperf.	4	30	12	

6 The percentage is the ratio of double object constructions with regard to the sum of double and single object occurrences. The Fischer′s exact test examines the significance of the association between the aspectual character of the verb and the syntactic construction (double vs. single object); the significant level of 0.05 is used in this paper

Some verbs with high double object occurrence are not testable because there are not enough (if any) aspectual counterparts for proper testing. The results show that except for two verb pairs

(2) *dát* – *dávat*
give-Perf. – give-Imperf.

(3) *předložit* – *předkládat*
submit-Perf. – submit-Imperf.

there is a positive correlation between perfectivity and double object syntax. However, a statistical testing, namely the Fischer's exact test, reveals that there are only two pairs whose differences between occurrences of the double object constructions with regard to aspect are statistically significant (with 0.05 significant level):

(4) *doporučit* – *doporučovat*
recommend-Perf. – recommend-Imperf.

(5) *nabídnout* – *nabízet*
offer-Perf. – offer-Imperf.

What does the result show? First, it is obvious that the application of statistical methods makes the observation more reliable. What looks at the first sight as a confirmation of the hypothesis (at least as a tendency – nine of eleven perfective verbs have the tendency to occur more often in a double object construction than their aspectual counterparts) was radically changed by statistical testing of the results. The test shows that except for verb pairs (4) and (5) we cannot reject the null hypothesis. That means that the distribution of the double object construction is independent of the aspect at the 5% probability level. Consequently, the TH prediction of the relationship between two high transitivity parameters (presence of the indirect object and the perfective verb) is false.

However, the statistically significant differences in verb pairs (3) and (4) should be explained. Of course, only further testing using other language material will lead to more reliable conclusions. But some preliminary explanation can be offered now: perfective verbs in pairs (4) and (5) belong to the verbs with the highest double object syntax occurrences (see table 1), so it seems reasonable to consider the relationship between the hypothesis and the relative frequency of double object constructions. However, the perfective verb with the highest double object syntax in the list

(6) *předat*
hand over-Perf.

is ranked among the verb pairs with no statistically significant differences of distribution of double object construction with regard to aspect. A small number of occurrences however of imperfective verb form slightly detracts from the validity of the interpretation of the statistical test in this case. In other words, the slight change of distribution of imperfective verb leads to the different result of the statistical test.

On the other hand, the perfective verb with the fourth highest relative frequency of double object construction

(7) *poskytnout*
allow-Perf.

is ranked among the pair with the p-value on the very border of the significant level; although one does not reject the null hypothesis, the rejection is much weaker than in other cases.

These facts make it possible to consider a relationship between frequency and hypothesis validity and to formulate a new working hypothesis: the perfectivity correlates with double object verbs which have *the high relative frequency of double object constructions*. This working hypothesis has to be tested on other language material, of course. However, the database based on the PDT makes it possible to formulate a tentative test: the verbs which have high relative frequency of double object syntax (more than 50% of all occurrences of a given verb) but low overall frequency (less than 40 occurrences) can be used for the preliminary testing.

There are two pairs of testable verbs in the database:

(9) *zaslat* – *zasílat*
send-Perf. – send-Imperf.

(10) *dovolit* – *dovolovat*
permit-Perf. – permit-Imperf.

In each pair one of the verbs is used frequently in double object syntax (more than 50% of all occurrences of a given verb), as table 4 shows. We can see that there is a positive correlation between perfectivity and double object syntax and the statistical testing also reveals that there are significant differences between occurrences of double object constructions and aspect (at the 0.05 significant level). Therefore, the working hypothesis is not false.

Table 4: The occurrences of verbs with low overall frequency[7]

verb	aspect	ditransitive	monotransitive	percentage of ditransitive syntax	Fischer's test p-value
zaslat 'send'	perf.	16	4	80	**0.006**
zasílat 'send'	imperf.	3	8	27	
dovolit 'permit'	perf.	11	5	69	**0.004**
dovolovat 'permit'	imperf.	4	16	20	

5. Conclusion

The present study reveals that there is no statistically significant correlation between the two transitivity parameters, viz. the presence of an indirect object in a sentence and perfectivity. The results are similar as those found in Newman & Rice (2006) and Olsen & Macfarland (1996), who also revealed problems with some TH's predictions. A preliminary reinterpretation of the original hypothesis was offered. This reinterpretation also takes into

7 Occurrences of verbs with the high relative frequency of double object syntax (more than 50% of all occurrences of a given verb) but low overall frequency (less than 40 occurrences).

account the frequency of observed verbs; it predicts that the perfectivity correlates with the double object verbs which have the high relative frequency of double object constructions.

References

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *The Longman grammar of spoken and written English*. London: Longman.

Bybee, Joan & Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In J. Bybee & P. Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*. Amsterdam, Philadelphia: John Benjamins. 1–24.

Hajič, Jan, Jarmila Panevová & Eva Hajičová. 2006. *Prague Dependency Treebank 2.0*. Philadelphia: Linguistic Data Consortium.

Hopper, Paul & Sandra Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56: 251–299.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2007. Ditransitive constructions: a typological overview. Typescript, Max Planck Institute, Leipzig.

Mukherjee, Joybrato. 2005. *English Ditransitive Verbs*. Amsterdam, New York: Rodopi.

Naess, Åshild. 2007. *Prototypical Transitivity*. Amsterdam, Philadelphia: John Benjamins

Newman, John & Sally Rice. 2006. Transitivity schemas of English EAT and DRINK in the BNC. In S. Th. Gries & A. Stefanowitsch (eds.) *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter. 225–260.

Olsen, Mari Broman & Talke Macfarland. 1996. Where's Transitivity? Paper presented at the *Seventh Annual Formal Linguistic Society of Mid-america* conference, May 17–19 1996, The Ohio State University.

Popper, Karl R. 1951. *The Logic of Scientific Discovery*. London, New York: Routledge.

Thompson, Sandra & Paul Hopper. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. In J. Bybee & P. Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*. Amsterdam, Philadelphia: John Benjamins. 27–56.