

Pitfalls of the Transitivity Hypothesis: Transitivity in Conversation and Written Language in Czech¹

Radek Čech²

Petr Pajas³

Abstract: *The aim of the article is to test empirically predictions formulated in the Transitivity Hypothesis framework. Methodological problems of the original approach are discussed and some solutions are offered. For the testing of the hypotheses two corpora of Czech were used (Prague Spoken Corpus and Prague Dependency Treebank). The results question both the predicted impact of the language form on transitivity and, more importantly, the concept of the Transitivity Hypothesis in general.*

1. Introduction

Transitivity is a central phenomenon in the structure of human language and it is usually considered to be a universal property of the language. Traditionally (cf. Tsunoda 2005) transitivity refers to the sentence property which is determined by the presence (or absence) of an object. Thus, clauses containing an object are assigned as transitive

(1) *John hit Paul*

while clauses with no object are assigned as intransitive

(2) *Mary sleeps.*

From the semantic point of view, a transitive clause expresses an activity/action which goes from the subject to the object.

However, among linguists there is not any consensus about the character of transitivity, although the term is often used in a way which takes its content for granted (cf. Naess 2007). Moreover, an absolute majority of linguistic analyses (generally) have not used an empirical science methodology in Popper's (1959) sense. This means that an explanation of linguistic phenomena is not usually based on empirically testable hypotheses (cf. Köhler 2005, Sampson 2001a, 2005). Consequently, there is a host of approaches to transitivity and no manner how to intersubjectively single out, by means of empirical tests, the best one among them.

In this paper, neither particular approaches to transitivity will be compared nor another transitivity theory will be proposed. Since we are convinced that language should be observed by methods which are common in empirical sciences (Čech 2007), we decided to follow an approach to transitivity which enables empirical testing. More precisely, we consider the Transitivity Hypothesis (hereinafter TH), originally formulated by Hopper and Thompson (1980). TH predicts the universal properties of language and is ranked among influential theories of transitivity, especially in cognitive linguistics (cf. Geeraerts & Cuyckens 2007). Surprisingly enough, although predictions given by TH brought a new view on relationships

¹ Acknowledgements: R. Čech and P. Pajas were supported by GAČR (Czech Science Foundation) No. 405/08/P157 – *Components of transitivity analysis of Czech sentences (emergent grammar approach)* and by GA AV ČR (Information Society of Grant Agency of Academy of Sciences of the Czech Republic) No. 1ET101120503 — *Integration of language resources for information extraction from natural texts*, respectively.

² Radek Čech, University of Ostrava, Reální 5, Ostrava 701 03. Email: radek.cech@osu.cz

³ Charles University, Prague

among different grammatical and semantic facets of language (see Section 2), TH has been tested empirically very rarely (e.g., Olsen & Macfarland 1996, Otani 2008, Newman & Rice 2006, Čech 2009), and, surprisingly, the results have not always corroborated the predictions of TH.

This study is focused on the observation of the relationship between Transitivity⁴ and genre. Authors of the TH claim (Thompson & Hopper 2001) that language form (e.g., conversation) has a considerable impact on grammatical and semantic language properties which are ruled by Transitivity. However, this statement has important consequences for the TH in general, which have not been mentioned in Thomson & Hopper's original work (2001) and which undermine, at least partly, the concept of the TH (see Section 3). Moreover, the testing of the impact of the genre on Transitivity is based on statistical analyses and interpretations which are methodologically very problematic (cf. Altmann & Lehfeldt 2004).

In this article, we start by summarizing the main features of TH in Section 2. In Section 3 the methodological problems of the approach (Thompson & Hopper 2001) are pointed out and solutions to these problems are offered. Results of the hypotheses testing in Czech are presented in Section 4. A summary is presented in Section 5.

2. The Transitivity Hypothesis

Hopper & Thompson (1980) consider transitivity as a crucial relationship in language which has a number of universally predictable consequences in grammar. Transitivity is not viewed in a traditional sense, which means that the presence (or absence) of the object in the sentence is the only parameter distinguishing between transitive (or intransitive) clauses, but Transitivity is regarded as a continuum: it is a matter of the grammar of the entire clause and it “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); see Table 1. A value of Transitivity of the sentence is given by the presence of high Transitivity features, so, the sentence

(3) *Susan left*

is more Transitive than the sentence

(4) *Jerry likes beer*

because sentence (3) has more high Transitivity features (Kinesis: action; Aspect: telic; Punctuality: punctual; Volitionality: volitional) than sentence (4) (Participants: two) (ibid. p. 254).

Table 1: Transitivity parameters

	parameter	high Transitivity feature	low Transitivity feature
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

⁴ The term Transitivity (with capital T) refers to Hopper & Thompson's (1980) concept of transitivity, see Section 2.

	parameter	high Transitivity feature	low Transitivity feature
F	AFFIRMATION	affirmative	negative
G	MODE	realis	irrealis
H	AGENCY	Agent high in potency	Agent low in potency
I	AFFECTEDNESS of Object	Object totally affected	Object not affected
J	INDIVIDUATION of Object	Object highly individuated	Object non-individuated

The most important aspect of the TH lies in the prediction hypothesizing 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, p. 254).

3. The Transitivity Hypothesis with regard to the language form (spoken vs. written)

It is well known that genre importantly affects grammar of the text (cf. Biber 1999). In Thompson & Hopper (2001), the impact of the language form (i.e., conversation) on Transitivity is analysed. Concretely, the authors claim that conversation is very low in Transitivity because of its nature: “the low Transitivity in our conversational data is to a considerable extent determined by the kinds of things we are doing when we talk with friends and acquaintances. We do not seem to talk much about events, let alone actions (as Hopper 1991, 1997a has also shown), but rather, our talk is mostly about ‘how things are from our perspective’. Our data show that we describe states, reveal our attitudes, ascribe properties to people and situations, and give our assessments of situations and behavior” (2001, p. 53).

First, the distributions of clauses with one participant, on the one hand, and clauses with two or more participants, on the other, were observed for corroboration of this argument. The authors have predicted that if Transitivity is low in conversation, the majority of clauses turn out to have one participant. At the first sight, presented results seem to confirm the prediction: 73% of one participant clauses and 27% of two or more participant clauses were detected in the dataset, consisting of face-to-face multi-party conversations among friends and family members. Nevertheless, what does it actually mean when one says that something is ‘low’ or ‘high’ without an explicit scale factor? In other words, how much percentage of one participant clauses is ‘enough’ to say that Transitivity is low?

Obviously, there is an assumption that Transitivity is high in constructed sentences and spoken or written narratives (cf. Thompson & Hopper 2001, p. 27, 52); but the authors have not formulated explicitly that Transitivity is low *in comparison* to written language (or particular genre). In other words, the statement, considering the relationship between the language form and Transitivity, does not have a form of a testable hypothesis and, consequently, the interpretation of the results is questionable. But this insufficiency could be set right easily, if one hypothesizes, for instance, that *the ratio of one participant clauses, in comparison to two or more participant clauses, is higher in conversation than in written language*. The testing of this hypothesis in Czech and the results are presented in Section 4.

Low Transitivity in conversation is also manifested by the character of two participant clauses, according to the authors. If Transitivity is considered to be a scalar value, consisting of 0-10 high Transitive parameters, and the number of participants is just one of the ten parameters, also two or more participant clauses can be low Transitive, in fact (see sentence (4)

above). Thompson & Hopper claim that the observation of two or more participant clauses in the dataset (conversation) confirm their prediction (see Table 2). The percentage expresses the ratio of low Transitivity features in two participant clauses in a conversation.

Table 2. Low transitivity of two-participant clauses (based on Thompson & Hopper 2001, p. 37).

Kinesis: Non-action	86%
Aspect: Atelic	86%
Punctuality: Non-punctual	98%
Affectedness: Non-affected Object	84%
Mode: Non-irrealis	70%
Individuation: Non-individuated Object	55%
Volitionality: Non-volitional Agent	50%
Agency: Potent Agent	97%

However, the interpretation of the findings presented in Table 2 is problematic; (1) there are unclear methodological aspects in the approach, (2) the findings have important consequences for the concept of the TH in general. As for the former, the statements considering the low Transitivity in two participant clauses lack the form of hypothesis. Moreover, Hopper and Thompson have not given clear and unambiguous criteria for distinguishing some semantic parameters. It is probably assumed that the evaluation of these criteria, such as Volitionality or Affectedness of object, are obvious, although the observation of semantic properties of language fumbles with many problems (cf. Sampson 2001b). Therefore, we suggest reformulating the original statement to a testable hypothesis, and, at least in the beginning, observing parameters which can be unambiguously determined. For instance: *in conversation, there are more imperfective predicates in two or more participant clauses than in written language*. For more detail see Section 4.

As for the consequences of the findings for the concept of the TH, it is necessary to mention that the TH predicts systematic co-variation of high (or low) Transitivity features: “whenever two values of the Transitivity components are necessarily present (...) they will agree in being either both high or both low in value” (1980, p. 254). It logically means that “the opposite type of correlation will not be found, where a high-Transitivity feature systematically co-varies with a low-Transitivity feature in the same clause” (Hopper & Thompson 1980, p. 255). However, the results in Table 2 indicate co-variation of opposite features, which is in direct contradiction with the prediction of the TH. More concretely, two-participants are treated as high Transitivity feature (see Table 1) and two participants should not correlate with low Transitivity features as Non-action, Atelic, Non-punctual, and so on. However, the findings presented in Table 2 shows co-variation between two-participants, on the one hand, and Non action, Atelic, Non-punctual and another low Transitivity features, on the other.

The consequence is obvious: if the predictions given by the TH are rejected, the whole concept of the TH is jeopardized, at least in conversation (but originally the TH should be valid universally).

4. Transitivity in spoken and written Czech

In the previous section some hypotheses which enable us to empirically test the statements considering the relationship between language form and Transitivity were proposed. Before the results of analyses are presented (Section 4.2), the material and methods are described in the next section.

4.1 Language material and methods

For the testing of the hypotheses two corpora of Czech were used: the Prague Spoken Corpus (Čermák 2007) and the Prague Dependency Treebank 2.0 (Hajič et al. 2006). The Prague Spoken Corpus (hereinafter PSC) captures real spoken Czech and it covers the four socio-linguistic variables in balanced proportions: the speaker's gender, age, education and type of speech. Only informal conversations were used in this study because they are the most similar to the language material in Thompson & Hopper's paper (2001): in the PSC, informal conversations are spontaneous dialogues between speakers who know each other and the topics of their conversations were not determined beforehand. The Prague Dependency Treebank (hereinafter PDT) consists of articles from newspapers and journals.⁵

Both corpora are tagged morphologically and annotated on a syntactic level. Although syntactic annotations in both corpora are not identical (e.g., the PDT's annotation is much more complex than the PSC's annotation, which assigned only valency frames), they are both based on the dependency formalism. The differences between the PSC and the PDT's annotation do not have a significant impact on the results and their interpretation, according to us, because the two corpora are not compared directly; only differences between distributions of observed characteristics *within* each corpus are a matter of testing.

In each of our experiments, we test a pair of properties, assuming a null hypothesis stating that the two properties are independent. Such a test can either reject the null hypothesis in favour of the alternative hypothesis that there is some correlation between the observed properties, or fail to reject the null hypothesis. The usual statistical method for performing this type of tests is the Pearson's chi-square test, where the probability of the observed data is computed assuming the χ^2 distribution expected by the null hypothesis; the lower the computed *p*-value, the less likely it is that we observe the data assuming the null hypothesis. The level of significance for the *p*-value will be set to 0.05 (5%). Since it is impossible to test statistically the results obtained from whole corpora (statistical tests were designed for a much smaller amount of data; with sample sizes of tens of thousands almost all null hypotheses are rejected), subcorpora of the PSC and the PDT were used; each subcorpus consisting of approximately the first 25,000 tokens of the original corpus.⁶

4.2 The results

In this section the results of two hypotheses testing are presented. Firstly, the differences between distributions of one participant and two or more participant clauses will be observed. Secondly, the distribution of perfective and imperfective predicates in two or more participant clauses in conversation and written language will be scrutinized.

If the conversation is low in Transitivity (in comparison to written language), there should be significant differences between distributions of one participant clauses in the PSC

⁵ For more details see web page: <http://ufal.mff.cuni.cz/pdt2.0/>

⁶ The queries that were used to extract quantitative data are available at the web page: <http://ufal.mff.cuni.cz/~pajas/papers/TH.html>

and the PDT. In other words, the ratio of one participant clauses should be higher in conversation than in written language. As Table 3 shows, our findings do not corroborate this prediction. On the contrary, the ratio of one participant clauses is higher in the PDT, but the difference between both distributions is not statistically significant (the observed data do not reject the null hypothesis of independence).

Table 3. The distribution of one and two or more participant clauses in the PSC and the PDT

	1 participant clauses	2 or more participant clauses	% 1 participant clauses
PSC	292	2054	12.5%
PDT	230	1348	14.6%

$$\chi_1^2 = 3.71, P = 0.0542$$

Next, the distribution of perfective (and imperfective) predicates in two or more participant clauses is observed. The Aspect, parameter C, was chosen for the testing because (1) the aspect is formally well distinguishable in Czech; (2) the aspect is an independent parameter (cf. Olsen & Macfarland 1996), and (3) in Karlík (2000), the TH is used for the explanation of the relationship between the number of participants of the sentence and aspect in Czech⁷. In conformity with Thompson & Hopper's prediction, the ratio of imperfective predicates is higher in the PSC and the difference between distributions in the PSC and PDT is statistically significant, as Table 4 shows.

Table 4. The distribution of imperfective and perfective predicates in two or more participant clauses in the PSC and the PDT

	imperfective predicates	perfective predicates	% imperfective predicates
PSC	1715	324	84.1%
PDT	887	460	65.9%

$$\chi_1^2 = 150, P \sim 0$$

At the first sight, the results confirm Thompson & Hopper's statement about the low Transitivity in conversation — the co-variation between the number of participant, parameter A, and the Aspect, parameter C, is stronger in the PSC than in the PDT. But if the distributions of imperfective and perfective predicates in one participant clauses in the PSC and the PDT are observed, one can see (Table 5) that the results are similar to the results considering two or more participant clauses; the ratio of imperfective predicates is higher in the PSC and the difference between distributions in the PSC and PDT is statistically significant.

⁷ Of course, it can be objected that the perfectivity-imperfectivity dichotomy does not correspond to telic-atelic dichotomy. We decided to use the perfectivity-imperfectivity dichotomy mainly for the sake of clarity.

Table 5. The distribution of imperfective and perfective predicates in one participant clauses in the PSC and the PDT

	imperfective predicates	perfective predicates	% imperfective predicates
PSC	245	44	84.8%
PDT	158	72	68.7%

$$\chi_1^2 = 19.08, P = 0.00001$$

Moreover, if the distributions of imperfective and perfective predicates with regard to the number of participants within each corpus are observed, no statistically significant differences are found (Table 6 and 7). This means that there is no statistically significant co-variation between the number of participants and the Aspect in both corpora. In other words, the TH's prediction considering this relationship does not stand regardless to the language form. On the other hand, although the language form influences the use of imperfective predicates, it does so regardless of the number of predicates, at least in observed datasets.

Table 6. The distribution of imperfective and perfective predicates in one and two or more participant clauses in the PSC

	imperfective predicates	perfective predicates	% imperfective predicates
1 participant clauses	245	44	84.8%
2 or more participant clauses	1715	324	84.1%

$$\chi_1^2 = 0.08, P = 0.772$$

Table 7. The distribution of imperfective and perfective predicates in one and two or more participant clauses in the PSC

	imperfective predicates	perfective predicates	% imperfective predicates
1 participant clauses	158	72	68.7%
2 or more participant clauses	887	460	65.9%

$$\chi_1^2 = 0.71, P = 0.399$$

5. Conclusion

The study reveals some methodological deficiencies in Thompson & Hopper's (2001) approach. The empirical testing of the hypotheses led to the rejection of two Thompson & Hopper's predictions. First, it was shown that there are no statistically significant differences

between distributions of one participant and two or more participant clauses in spoken and written language. Second, the TH prediction hypothesizing the relationships between the number of participants and the Aspect was rejected regardless of the language form. As for the latter, this finding questions the concept of the TH in general and it coincides with studies that also rejected some of the TH predictions, at least partly (Olsen & Macfarland 1996, Newman & Rice 2006, Čech 2009). However, an appropriate evaluation of the TH needs scrutinizing TH predictions in more languages.

References

- Altmann, G. - Lehfeldt, W.** (2004). Book review. (Bybee, J. - Hopper, P. (eds.) (2001): *Frequency and the Emergence of Linguistic Structure*. Amsterdam, Philadelphia: John Benjamins). *Journal of Quantitative Linguistics* 11, 275–304.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E.** (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Čech, R.** (2007). Language System - Linguistics as an Empirical Science. *Sapostavitelno ezikoznanie* 32, 42-49.
- Čech, R.** (2009). Testing of the Transitivity Hypothesis: Double Object Verbs and Aspect in Czech. In Dočekal, M., Ziková, M. (eds.), *Czech in Formal Grammar*. München: Lincom, 29-38.
- Čermák, F.** (ed.) (2007). *Frekvenční slovník mluvené češtiny*. Praha: Karolinum.
- Geeraerts, D., Cuyckens, H.** (eds.) (2007). *The Oxford Handbook of Cognitive Linguistics*. Oxford.
- Hajič, J., Panevová, J., Hajičová, E., Sgall, P., Pajas, P., Štěpánek, J., Havelka, J., Mikulová, M.** (2006). *Prague Dependency Treebank 2.0*. Philadelphia: Linguistic Data Consortium.
- Hopper, P., Thompson, S.** (1980). Transitivity in Grammar and Discourse. *Language* 56, 251-299.
- Hopper, P.** (1991). Dispersed verbal predicates in vernacular writing. *BLS* 17, 402-413.
- Hopper, P.** (1997a). Discourse and the category 'verb' in English. *Language and Communication* 17, 93-102.
- Karlík, P.** (2000). Hypotéza modifikované valenční teorie. *Slovo a slovesnost* 61, 170 -189.
- Köhler, R.** (2005). Synergetic linguistics. In: Köhler, R., Altmann, G., Piotrowski, R.G. (eds), *Handbook of Quantitative Linguistics*. Berlin: de Gruyter, 760-775.
- Naess, Å.** (2007). *Prototypical Transitivity*. Amsterdam, Philadelphia: John Benjamins.
- Newman, J., Rice, S.** (2006). Transitivity schemas of English EAT and DRINK in the BNC. In: Gries, S.T., Stefanowitsch, A. (eds.), *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*. Berlin, New York: Mouton de Gruyter, 225–260.
- Olsen, M.B., Macfarland, T.** (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.
- Otani, N.** (2008). The Transitivity Hypothesis and Object Case-marking in Japanese: On the Transitivity of oyogu (Swim) in Japanese,"Workshop on East Asian Linguistics, Santa Barbara, February 23, 2008.
- Popper, K.R.** (1959). *The Logic of Scientific Discovery*. London: Hutchinson & Co Routledge.
- Sampson, G.** (2001a). *Empirical Linguistics*. London/New York: Continuum.
- Sampson, G.** (2001b). Meaning and the Limits of Science. In *Empirical Linguistics*. London/New York: Continuum, 180-107.

- Sampson, G.** (2005). Quantifying the Shift Towards Empirical Methods. *International Journal of Corpus Linguistics* 10, 15-36.
- Thompson, S., Hopper, P.** (2001), Transitivity, clause structure, and argument structure: Evidence from conversation. In: Bybee, J., Hopper, P. (eds.): *Frequency and the Emergence of Linguistic Structure*. Amsterdam, Philadelphia: John Benjamins, 27– 56.
- Tsunoda, T.** (2005). Transitivity. In: Brown, K. (ed.): *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon, 4670-4677.