

Objectivity and Reproducibility of Proppian Narrative Annotations

Rens Bod¹, Bernhard Fisseni^{1,2,3,4}, Aadil Kurji¹, Benedikt Löwe^{1,3,4,5}

¹ Institute for Logic, Language and Computation, Universiteit van Amsterdam
Postbus 94242, 1090 GE Amsterdam, The Netherlands
{l.w.m.bod,b.loewe}@uva.nl, aadil.kurji@student.uva.nl

² Fakultät für Geisteswissenschaften, Universität Duisburg-Essen
Universitätsstraße 12, 45117 Essen, Germany
bernhard.fisseni@uni-due.de

³ Isaac Newton Institute for Mathematical Sciences
20 Clarkson Road, Cambridge CB3 0EH, United Kingdom

⁴ Fachbereich Mathematik, Universität Hamburg
Bundesstraße 55, 20146 Hamburg, Germany

⁵ Corpus Christi College, University of Cambridge
Cambridge CB2 1RH, United Kingdom

Abstract

A formal narrative representation is a procedure assigning a formal description to a natural language narrative. One of the goals of the *computational models of narrative* community is to understand this procedure better in order to automatize it. A formal framework for automatization should allow for objective and reproducible representations. In this paper, we present empirical work focussing on objectivity and reproducibility of the formal framework by Vladimir Propp (1928). The experiments consider Propp’s formalization of Russian fairy tales and formalizations done by test subjects in the same formal framework; the data show that some features of Propp’s system such as the assignment of the characters to the *dramatis personae* and some of the functions are not easy to reproduce.

1. Introduction & Motivation

The formal study of narratives goes back to the Russian structuralist school, paradigmatically represented by Vladimir Propp’s 1928 study *Morphology of the Folktale* (Propp, 1958). Researchers in the field of *computational models of narrative* have developed the general Proppian methodology into formal and computational frameworks for the analysis, automated understanding and generation of narratives.¹

In recent years, there has been an increased interest in the methodological and conceptual issues involved. The enterprise of representing a narrative by a formal structure that can then be used in computational application rests on a number of assumptions:

Assumption E. (*Existence of a structural core*) There is a structural core of narratives; or several, depending on which part of the structure we are interested in.

Assumption O. (*Objectivity of the structural core*) Given a narrative, there is an interpersonal agreement what its structural core is; possibly after some agreement of what part of the structure should be represented.

A formal framework Λ for representing narratives consists of a formal language \mathcal{L}_Λ , a class of mathematical structures \mathcal{M}_Λ , and a description of a procedure (called *formalization* in (Löwe, 2011)) of assigning to each natural language narrative N a structure $\Sigma_\Lambda(N) \in \mathcal{M}_\Lambda$. Note that this procedure

is not a function in the mathematical sense, but an activity by expert formalizers who follow given guidelines.

In this paper, we explore the validity of Assumption **O**: in particular, we are investigating the following property of formal frameworks Λ :

Property **Obj(Λ).** Sufficiently trained human formalizers, given the same narrative N will produce the same structure $\Sigma_\Lambda(N)$.

Property **Obj** is an important (and arguably necessary) feature of a formal framework Λ if it is supposed to be the basis of an automatized system. The existence or non-existence of formal frameworks Λ with property **Obj**(Λ) is closely related to Assumption **O**. In (Bod et al., 2011), we described the investigation of **Obj**(Λ) as a natural analogue of the study of annotator agreement in corpus linguistics and computational linguistics: whereas typical annotation tasks involve annotation of sentences or discourses (e.g., (Marcus et al., 1993; Brants, 2000; Passonneau et al., 2006)), the formalization or annotation of a narrative is at the next level of complexity. At the sentence or discourse level, inter-annotator agreement has been studied (e.g. (Carletta et al., 1997; Marcu et al., 1999)), but no such analysis has been done for the formalization of narratives, not even for the oldest and best-known formal approach, Propp’s *Morphology of the folktale*, first published in 1928.

We focus on this formal framework, not because it is a particularly good candidate for a framework close to the *stable structural core*, but due to its prominent place in the history of formal representations of narratives. In § 2., we describe the Proppian formal framework and discuss two empirical studies pertaining to it, referred to as **Propp I** and **Propp II**, performed at the *Universiteit van Amsterdam*; in § 3., we discuss the results and future work.

¹Lehnert’s Plot Units, Rumelhart’s Story Grammars, Schank’s Thematic Organization Points (TOPs), Dyer’s Thematic Abstraction Units (TAUs), or Turner’s Planning Advice Themes (PATs) are some examples; cf. (Lehnert, 1981; Rumelhart, 1980; Schank, 1982; Dyer, 1983; Turner, 1994).

Ivanko								
Test subject	H	V	P	PF	Di	Do	MH	FH
1	Ivanko	Devils			Peasant			
2	Bearlet	Bear/Devil	Bearlet/Wife	Peasant	Peasant			
3	Ivanko	Thieves/Dogs/Devil	Wife	Peasant	Peasant	Grandfather	Horse	
4	Ivanko/Mother	Devil/Peasant			Peasant			
5	Ivanko	Father	Father's Satisfaction	Grandfather	Father	Little Devil	Horse	
6	Ivanko	Devil			Peasant			
7	Ivanko	Devil	Bear	Ivanko/Wife		Peasant	Devil	
8	Ivanko	Devil			Peasant		Horse	
9	Bearlet	Father	Father, Money		Father	Devil	Hare	

Semyons								
Test subject	H	V	P	PF	Di	Do	MH	FH
1	Semyons		Elena		Tsar			
2	Semyons	Tsar	Elena		7			
3	Semyons	7th Semyon	Elena	Tsar	Tsar	Kitten/Stone	7th Semyon	
4	7th Semyon	Elena's father	Elena	Elena's father	Tsar	Semyon Bros		
5	Semyons	Tsar	Elena	Elena's Father	Tsar	Tsar	Cat	
6	7th Semyon	Tsar	Elena	Elena's father	Tsar	Semyons		
7	Semyons		Elena	Tsar	Tsar	Ship		
8	7th Semyon		Elena		Tsar	6 Semyons		
9	7th Semyon	Tsar	Elena		Tsar	Semyons	Tsar	

Shabarsha								
Test subject	H	V	P	PF	Di	Do	MH	FH
1	Shabarsha			Gold	Little Devil/Grandad	Master		
2	Shabarsha/Little Devil	Shabarsha			Grandad	Grandad		
3	Shabarsha	Little Devil	Gold	Grandad	Grandad	Master	Cap	
4	Shabarsha	Little Devil	Gold	Grandad	Grandad	Master		
5	Shabarsha	Shabarsha	Gold	Grandad	Grandad	Master	Bear/Hare	Shabarsha
6	Shabarsha	Little Devil	Gold	Grandad	Master		Bear/Hare	
7	Shabarsha	Little Devil/Grandad	Gold		Master	Master	Twine	
8	Shabarsha	Little Boy	Gold	Grandad	Master		Bear/Hare	
9	Little Devil	Shabarsha	Peace	Grandad	Grandad			

Table 1: The assignment of the *dramatis personae* for the three folktales in **Propp I**.

2. Propp's formal system

2.1. Overview of Propp

Working with a corpus of 100 Russian folktales from Afanas'ev's collection *Narodnye Russkie Skazki*, Vladimir Propp developed a formal system to identify each folktale by short annotation strings consisting of symbols representing Proppian *functions* or *narratemes*. In the following, we give a description of the components of the Proppian system relevant for the experiments discussed in this paper. For more details, we refer the reader to (Propp, 1958).

Propp identified seven² *dramatis personae* representing roles the characters may play within the tales. They are: the hero (**H**), the villain (**V**), the princess (**P**), the princess's father (**PF**), the dispatcher (**Di**), the donor (**Do**), the (magical) helper (**MH**) and the false hero (**FH**) (Propp, 1958, § 3).

The actions of the *dramatis personae* are described by a set of thirty-one functions described in (Propp, 1958, § 3) by means of examples and more specified subfunctions. These functions are marked by symbols in the order of their occurrence in the folktale; the first seven functions, marked with lowercase Greek letters, are called *preliminary functions*: β Absentation; γ Interdiction; δ Violation, ϵ Reconnaissance, ξ Delivery, η Trickery, θ Complicity. The *preliminary functions* are not fully developed in (Propp, 1958) and are not included in Propp's own annotation strings. The main functions are: **A** Villainy, **a** Lack, **B** Mediation, **C** Be-

²One of these, the *Princess/Princess's Father*, can be split into two with a slightly difficult delineation. In our experiment, we presented the resulting list of eight *dramatis personae*.

ginning counteraction, \uparrow Departure, **D** First function of the Donor, **E** Hero's reaction, **F** Provision or receipt of magical agent, **G** Spatial transference between two kingdoms, **H** Struggle, **J** Branding, **I** Victory, **K** Liquidation, \downarrow Return, **Pr** Pursuit, **Rs** Rescue, **o** Unrecognized Arrival, **L** Unfounded Claims, **M** Difficult Task, **N** Solution, **Q** Recognition, **Ex** Exposure, **T** Transfiguration, **U** Punishment, **W** Wedding. These functions, instantiated by *subfunctions* marked by superscripts, occur in strict sequential order, i.e., functions have to occur in the folktale in the order they are given in the list above. In the full Proppian system, there are a few specific ways to break strict sequentiality (Propp, 1958, § IX.A): The most important one is that some folktales contain a series of individual tale units, called *moves*. Examples are *trebling*, the triple repetition of moves within the tale, and moves in which a magical agent is obtained in the first move but only used in the second move of the tale. None of the tales we used had *moves* (according to Propp's own annotations), so we did not include this option in our experiment.

2.2. Description of Propp I

Test subjects were trained in the Proppian framework and then asked to annotate four of the folktales formalized in (Propp, 1958). We used the folktales *The Seven Semyons*, *147, Shabarsha*, *151*, and *Ivan the Bear's Son*, *152*; in the following, we refer to these folktales as *Semyons*, *Shabarsha*, and *Ivanko*.³ We chose tales that were available in English translation, and in Propp's annotation had no *moves*

³In **Propp I**, we also used the folktale *The Enchanted Princess*, but it was too long and omitted in **Propp II**. Due to an

(i.e., retained strict sequential ordering) and used few functions (*Ivanko* uses eight functions, *Shabarsha* six). An annotation of in **Propp I** consisted of (1) the assignment of story characters to the *dramatis personae*, and (2) a list of the functions (group 1) or the functions with corresponding subfunctions (group 2) occurring in the folktale.

Procedure. We had nine test subjects, all students of the *Universiteit van Amsterdam*, and all with native or near-native competence of English. We split them into two groups: Test subjects 1 to 5 were group 1 (no subfunction marking) and test subjects 6 to 9 were group 2 (subfunction marking). Test subjects were instructed that the experiment would last three hours and received a moderate financial compensation for participation.

The experiment started with a 45-minute introduction to Propp’s system given by a native speaker of English supported by a projector presentation explaining the relevant fragment of Propp’s system. Only a selection of the subfunctions was included (labelled “examples” for group 1 and “subfunctions” for group 2). We analyzed a simple example story, of our own design, as an illustration. A condensed version of the *dramatis personae* and functions was distributed as a leaflet for use during the annotation.

Results. Propp’s annotation for *Shabarsha* was $A^8 B^4 C \uparrow H^2 I^2 K^1 \downarrow$; his annotation for *Ivanko* was $A^9 \uparrow H^2 I^2 K^1 \downarrow$. These consist of the function strings alone and do not include the *preliminary functions*.⁴

We give the results of the assignments of *dramatis personae* in Table 1. The results indicate that the test subjects did not fully understand the Proppian scheme; note in particular the variation in the three main *dramatis personae*, **H**, **V**, and **P** (see below for a methodological remark).

The annotation strings vary widely and are given in Table 2 (subfunctions are marked by superscripts, with a missing subfunction marked by \varnothing). Since no two strings are the same, comparison would have to be per function; calcula-

oversight, we worked with version 147 of *Semyons* while Propp annotated version 145. This makes it impossible to compare our results to Propp’s original annotation, but it does not invalidate the discussion of inter-annotator agreement of our test subjects. We used the translations of Gutermann (Afanas’ev, 1973) for *Semyons* and *Ivanko*, and the translation of Cook (Afanas’ev, 1985) for *Shabarsha*.

In *Semyons*, seven orphans meet the Tsar and pledge to work hard in their professions. The seventh becomes a thief and, with the help of his brothers and their respective talents, journeys to capture Elena the fair as a bride for the Tsar. In *Ivanko*, *Ivanko* is born of a peasant woman and her kidnapper, a bear. After returning to human society, he causes some damage and is sent to a lake in which devils dwell. Through a series of tricks, *Ivanko* gains all of the devils’ gold and the services of a little devil for a year. In *Shabarsha*, the protagonist *Shabarsha* takes a day off to earn some money for himself and his boss. He goes to a lake to catch fish, meets a little devil and threatens to evict all of the devils from the lake if they don’t pay rent. Through a series of tricks he acquires all of their wealth.

⁴Therefore, we do not take the *preliminary functions* into account for comparison between Propp’s original strings and the strings produced by the test subjects. For the sake of completeness, we also list the Propp string for *Semyons* (version 145, cf. fn. 3): $a^1 B^2 C \uparrow F^2 G^1 K^2 \downarrow$.

tions of statistics per function are not useful because of the variation in the assignment of *dramatis personae* and the small amount of data. The strings are longer than Propp’s strings (compare an average of 14.2, 13.2, and 12.8 functions with the Propp’s of 6 and 8 for *Ivanko* and *Shabarsha*, respectively).

Methodological Conclusion. Four out of nine test subjects reported that the example story from the presentation was considerably simpler than the folktales.

The variation in the assignment of characters to *dramatis personae* suggests that the description of the *dramatis personae* was not precise enough. For instance, our description of the hero used the words “who is good”. Arguably, *Shabarsha*’s behaviour in *Shabarsha* cannot be described as “good”, which caused some of the variation in the assignment of the hero.⁵

A number of functions are consistently annotated which do not show up in Propp’s own annotations. On the other hand, we see that some of Propp’s functions show up in all or almost all annotations strings: e.g., \uparrow , **H**, **I**, **K** and \downarrow are reliably reproduced in the *Ivanko* annotation strings. However, since we do not know which events in the tale the annotators marked with these functions, we cannot be sure whether these are actual reproductions of Propp’s assignments.

2.3. Description of Propp II

The experiment **Propp II** was a modified version of **Propp I**, taking the problems discussed in § 2.1. into account. We used the same folktales as in **Propp I**. An annotation of a folktale in **Propp II** consisted of (1) a list of the functions occurring in the folktale, and (2) marked text passages for each of the functions that occurred.

The main changes to **Propp I** were: the test subjects were given the assignment of *dramatis personae*; subfunctions were not discussed at all; the example story was from Propp’s own corpus. It should be noted that Propp only recorded the annotation strings, so that his choice of *dramatis personae* was extrapolated from (Propp, 1958).⁶

Procedure. We had six test subjects, all students of the *Universiteit van Amsterdam*, and all with native or near-native competence of English. Test subjects were instructed that the experiment would last three hours and received a moderate financial compensation for participation.

The experiment started with a 45-minute introduction to Propp’s system given by a native speaker of English supported by a projector presentation explaining the relevant fragment of Propp’s system. We gave short descriptions of the *dramatis personae* roughly based on Propp’s original text and the the descriptions of the functions from Propp’s

⁵It is conceivable that the designator “devil” created a connotation in the original audience of the folktale producing a very different reading of *Shabarsha*’s behaviour that cannot be reproduced in contemporary test subjects due to a lack of cultural context and contemporary sympathy for harmless “little devils”.

⁶In *Ivanko*, we assigned *Ivanko* to **H** and the Little Devil and the Grandfather jointly to **V**; in *Semyons*, we assigned the seventh *Semyon* to **H**, Elena the Fair to **P**, and the Tsar to **Di**; finally, in *Shabarsha*, we assigned *Shabarsha* to **H** and the Little Devil and the Grandfather jointly to **V**.

Subject		Proppian functions for Ivanko				
Propp		A ⁹	↑	H ²	I ² K ¹ ↓	
1	β		↑	G H I	↓	
2	β γ δ ζ η θ	A a B C↑		G H I K ↓		U
3	β γ δ ζ η θ	A a B C↑D E		H I K ↓	N Ex	
4	β γ δ ε ζ η θ	a B C↑D		G H I K ↓	ZNQ	U
5	β γ δ ζ η θ	a B C↑D E F		JI K		
6	β ² γ ² δ ² ε ² ζη ¹ θ ¹	B ³ C↑		G ³ H ²	I ² K ¹ ↓	
7	β ¹ δ ² ε ²	a ⁵ B ² C↑D ¹ E ¹		H ²	I ² K ² ↓	U
8	β ¹ γ ²		↑	G ³ H ²	I ² ↓	W ⁶
9	β ¹ ζ θ ¹	a ⁵ B ² ↑D ¹ E ² F ² G ²		K ¹ ↓		

Subject		Proppian functions for Shabarsha				
Propp		A ⁸	B ⁴ C↑	H ² I ² K ¹ ↓		
1		a B C↑		G H K		
2	β εζη θ	A a B C↑		G H I K ↓		W
3	βγ εζη θ	A B C D EF		G H I K		
4	βγ εζ	a B C↑ E		I K Pr	QU	
5	βγ η	a B C D F		K ↓	QU	
6	γ ² δ η ¹ θ ¹	a ²		H ² I ² K ¹ ↓	N	UW ⁶
7	γ ²	a ² B ² C↑D ¹		F ¹ G ³ H ² I ² K ¹	o	W ⁶
8		a ² B ¹ C↑		G ³ H ² I ² K ¹		
9	η ² θ ¹	a ² B ²		H ²		W ⁶

Subject		Proppian functions for Semyons				
1	β	a B C↑	G	K ↓		
2	β γ δ ε ζ η θ	Aa B C↑	G H K ↓	PrRs		W
3	β γ δ ε ζ η θ	Aa	DEF G	K ↓	Pr	NQTU
4	β γ δ	a B C↑D	G	HIK ↓	PrRs	Q
5	β ζ η θ	a B C↑	F G	K ↓		
6	β ² γ ¹ δ ζ	a ¹ B ² ↑	G ²	K ¹ Pr	N	W ⁶
7		a ⁵ B ¹ C↑	F ³ G ³	K ¹ ↓		W ⁶
8	β ² γ ¹	a ¹ B ¹ C↑	G ³	K ² ↓		W ⁶
9	β ² γ ² δ ζ	a ⁵ B ² ↑	G ²	K ² ↓	Pr	N W ⁶

Table 2: The annotation strings for the three folktales in **Propp I** (cf. fn. 3).

Ivanko		Semyons		Shabarsha	
Test subject	Proppian Functions	Test subject	Proppian Functions	Test subject	Proppian Functions
Propp	A ↑ HIK↓			Propp	A BC↑HIK↓
1	β a ↑GHIK↓	1	aB ↑GK W	1	a HIK N
2	βγ ↑GHIK MN W	2	aB ↑ K↓ W	2	aB ↑ MN W
3	β B↑ HI U	3	β aB G oNW	3	a C↑HI M UW
4	β ↑ HI ↓ U	4	η a ↑G Pr W	4	a ↑HI MN
5	β aB↑ HI ↓	5	aB ↑ K↓PrRs W	5	a ↑H K
6	β aB↑ HIK↓ W	6	β aBC↑GK↓PrRs W	6	aBC HIK W

Table 3: The annotation strings for the three folktales in **Propp II** (cf. fn. 3).

text. We analyzed the folktale (*Ivan Popyalov, 135*) from the Propp corpus. Again the condensed version of the *dramatis personae* and functions was distributed as a leaflet to for use during the annotation. Test subjects were given an assignment of characters to the *dramatis personae* together with each folktale.

Results. We give the results of the function annotation in Table 3. The annotation strings are noticeably shorter than in **Propp I** (on average 6.8 functions per annotator, compared with 13.4 functions in **Propp I** and 6 and 8 functions in the original Propp strings for *Ivanko* and *Shabarsha*, respectively),⁷ and more similar to Propp’s original strings, but we still do not have matching strings among the test subjects.

It is again not possible to do a serious statistical analysis on the basis of six annotations; we therefore do a qualitative analysis instead. We say that a function occurs *stably* in **Propp II** if it is in at least four of the six annotations. We further distinguish *strong stability* when the marked text of the annotators overlaps, and *weak* otherwise. In *Ivanko*, β, I and ↓ were strongly stable and ↑ and H were weakly stable (of which ↑, H, I and ↓ are annotated by Propp); in *Shabarsha*, a and ↑ were strong stable and H and I were

weakly stable (of which ↑, H, and I were annotated by Propp); in *Semyons*, a, B, G, and W were strongly stable and ↑ and K were weakly stable. Note that in both *Ivanko* and *Shabarsha*, there is a strongly stable function not annotated by Propp (B and a, respectively).

3. Discussion & Future Work

The difference between **Propp I** and **Propp II** show that the assignment of the characters to the *dramatis personae* has an important effect on the assignment of the functions. Even with pre-assigned *dramatis personae*, there are marked differences between Propp’s and the test subjects’ annotations, and among the test subjects. Some of this effect can be explained by the vagueness of the description of Propp’s functions: as an illustration, we mention that subfunction 6 of W is listed as “Other form of compensation like a monetary reward”. This vague description fits in much more general situations than Propp apparently intended. Making these vague descriptions understandable for the test subjects may require considerably more time and training than we gave the test subjects in our experiments.

The detailed study of human annotations of Propp’s framework highlights weaknesses such as vague descriptions of *dramatis personae* and functions, and in general, points to some important obstacles for an automatization of the process of formalization in a computational setting.

⁷Most likely, a reason for the much longer strings in **Propp I** was the assignment of superfluous *dramatis personae* by the test subjects in that experiment.

In (Bod et al., 2011), we suggested to follow up the studies **Propp I** and **Propp II** with a large-scale inter-annotator study: the results of our experiments suggest that this is not worthwhile. Instead, we should distill the lessons learned from this Proppian case study into studies dealing with other formal representation systems, possibly designed and documented on the basis of the results of this study.

Acknowledgements

The research in this paper was funded by the *John Templeton Foundation* (JTF) via the project *What makes stories similar?* (grant id 20565) and the *Nederlandse Organisatie voor Wetenschappelijk Onderzoek* (NWO) via the projects *Integrating Cognition* in the VICI programme (DN 277-70-006) and *Dialogical Foundations of Semantics* in the ESF EuroCoRes programme LogICCC (LogICCC-FP004; DN 231-80-002; CN 2008/08314/GW). The second, and fourth author acknowledge the financial support and the kind hospitality of the *Isaac Newton Institute for Mathematical Sciences* (programme *Semantics & Syntax*). All authors would like to thank Ekaterina Abramova and Sanchit Saraf (Amsterdam) for their work in the early set-up phase of the research.

4. References

- Aleksandr Nikolayevich Afanas'ev. 1973. *Russian Fairy Tales*. Pantheon. Translation by Norbert Guterman from the collections of Aleksandr Afanas'ev. Folkloristic commentary by Roman Jakobson.
- Aleksandr Nikolayevich Afanas'ev. 1985. Shabarsha, translated by Kathleen Cook. In *The Three Kingdoms. Russian Fairy Tales From Alexander Afanasiev's Collection, illustrated by Alexander Kurkin*. Raduga Publisher, Moscow.
- Rens Bod, Benedikt Löwe, and Sanchit Saraf. 2011. How much do formal narrative annotations differ? A Proppian case study. In Charles Ess and Ruth Hagengruber, editors, *The computational turn: Past, presents, futures?*, Aarhus University, July 4–6, 2011, pages 242–245, Münster. MV-Wissenschaft.
- Thorsten Brants. 2000. Inter-annotator agreement for a German newspaper corpus. In *Proceedings Second International Conference on Language Resources and Evaluation LREC-2000*.
- Jean C. Carletta, Amy Isard, Steven Isard, Jacqueline C. Kowtko, Gwyneth Doherty-Sneddon, and Anne H. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31.
- Michael G. Dyer. 1983. *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Artificial Intelligence Series. MIT Press, Cambridge MA.
- Wendy G. Lehnert. 1981. Plot units and narrative summarization. *Cognitive Science*, 4:293–331.
- Benedikt Löwe. 2011. Methodological remarks about comparing formal frameworks for narratives. In Patrick Allo and Giuseppe Primiero, editors, *Third Workshop in the Philosophy of Information, Contactforum van de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten*, pages 10–28, Brussel. KVAB.
- Daniel Marcu, Magdalena Romera, and Estibaliz Amorortu. 1999. Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In Marilyn A. Walker, editor, *Towards Standards and Tools for Discourse Tagging. Proceedings of the Workshop*, pages 71–78.
- Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19:302–330.
- Rebecca Passonneau, Nizar Habash, and Owen Rambow. 2006. Inter-annotator agreement on a multilingual semantic annotation task. In *Proceedings LREC-2006*, pages 1951–1956.
- Vladimir Yakovlevich Propp. 1958. *Morphology of the folktale*. University of Texas Press, Austin, Texas, 2nd edition.
- David E. Rumelhart. 1980. On evaluating story grammars. *Cognitive Science*, 4:313–316.
- Roger C. Schank. 1982. *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press.
- Scott Turner. 1994. *The creative process. A computer model of storytelling*. Lawrence Erlbaum Associates, Hillsdale, NJ.