Korean Journal of English Language and Linguistics, Vol 22, December 2022, pp. 1368-1388
DOI: 10.15738/kjell.22..202212.1368

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

A Quantitative Analysis of the *Little Red Riding Hood* Types and Story Element-Function-Plot Relations*

Jungsik Park · Ho Han (Ajou University)

Received: October 06, 2022
Revised: December 12, 2022
Accepted: December 30, 2022

Jungsik Park (1st author)
Professor, Dept. of English Language and Literature, Ajou University
Email: jspark@ajou.ac.kr

Ho Han (corresponding author)
Professor, Dept. of English Language and Literature, Ajou University
Email: hhan@ajou.ac.kr

* This work was supported by the Ajou University research fund.

We would like to thank two anonymous reviewers for their valuable suggestions and comments. Of course, all errors are the authors' responsibility.

ABSTRACT

Park, Jungsik and Ho Han. 2022. A quantitative analysis of the *Little Red Riding Hood* types and story element-function-plot relations. *Korean Journal of English Language and Linguistics* 22, 1368-1388.

Tale types of "Little Red Riding Hood" have survived through oral transmission in various areas including Europe, Africa, and Asia and can even be traced back to 10th century in a written form. This research presents quantitative analyses on the folkloric landscape of tales of, or related to, what is best known as *Little Red Riding Hood* through the Aarne-Thompson-Uther (ATU) index, of which we analyzed ATU 333, ATU 123, and other unspecified types, based on logistic regression and decision tree. The quantitative analyses of the *Little Red Riding Hood* tale types indicate that ATU 123 alone has the specific story segments that are important to the formation of the tale type and that though diversified in story segments and other details, the three types shared the distinct plot sequence as an important feature. In addition, eight event descriptors and six character and setting descriptors are found to be meaningful factors in the formation of ATU 123. It can be further argued that the plot as an abstraction played a major role in the formation of the tales we have now. Also demonstrated in this paper is that researchers can yield substantial insights into the quantitative results while cross-checking them with qualitative analyses.

KEYWORDS

Little Red Riding Hood, logistic regression, feature importance, story element, plot, motif

1. Introduction

Recent years have witnessed a boom in the quantitative study of folktales, fueled by the adoption of computational methods and software for narrative analysis (Darányi et al. 1996, Roos and Heikkila 2009, Tehrani et al. 2016). The tale types of *Little Red Riding Hood*, which the current study deals with, have received much attention and have been measured quantitatively.

One reason that fueled quantitative research in the folktale studies is that folks telling tales are delivered from generation to generation and region to region, spawning numerous variants and versions, each leaving its own traces again in the milieu of folkloric networks. Medium-wise, folktales were orally transmitted and at a certain point of transmission, were recorded and written by folklorists, and then spread to another group orally or in written forms. Most well-known authors who recorded or revised the stories in Europe are Giambattista Basile (Italy), Charles Perrault (France), Joseph Jacobs (England) and Jacob and William Grimm (Germany), to name a few.

While the folklorists presented the tales editorially or authorially while borrowing from the tales told by common folks, their printed tales are just a tip of the folkloric iceberg, and the directions of transmission and variations of the same tales had been much diversified from the earliest to latest versions. Therefore, when studying a folktale as a whole, one may reach a very different conclusion, depending on the chosen samples of the variants poles apart from other research. For instance, Darnton (1984) criticized Bettelheim (1976) for not using the "original" version of *Little Red Riding Hood* (henceforth, LRRH) where, he believed, the story ends tragically—that is, the protagonist dies being swallowed by the wolf. Not to do injustice to Bettelheim's analysis, one needs to heed another LRRH story tradition according to which the protagonist is rescued by the hunter and takes her revenge on the wolf. Bettelheim (1976) chose to take up a strand of one tradition as Darnton (1984) did.

There are two major story traditions of LRRH types—according to the Aarne-Thompson-Uther (ATU) index system for the classification of folktales[1], ATU 333 which centers around Southern Europe and ATU 123 which covers Asia, the Middle East, Europe, and upper Africa. ATU 333 and ATU 123 are often cross-indexed as sibling types. Even between the two types, depending on variants and traditions, the victim becomes an innocuous little girl, audacious siblings, a group of animals, or the youngest of seven goat brothers. On the way to her mission, she often plucks up flowers, gathers pins, or in others simply goes straight to her grandmother. The Grimm brothers even decided to have two wolves as antagonists. In this folkloric milieu of storytelling and retelling, as Zipes (1993, p. 4) once pointed out, folklorists "frankly do not know what tale was the most 'common' at the time of transmission." Due to the limited extant resources, we cannot be sure of the extent of their originality and commonality when they were orally delivered from one person to another. Though it is impossible to figure out the commonest types in actu during the transmission, with hindsight we can better estimate quantitatively what features or story elements were most important for the tales to evolve into the LRRH types we have now.

Our research investigated what is now best known as, and represented by, the LRRH tale type, particularly ATU 333 and its sister type ATU 123, through logistic regression and decision tree. Amid various trajectories and possibilities, our research gravitated towards answering the following question: what are the most important story elements in forming the respective type of LRRH?

[1] The ATU index will be explained in Literature Review. For the time being, readers are recommended to take the ATU number is a sort of the name of a folktale group based on motif.

2. Literature Review

Folklorists have long been using various story element indexation systems, among which the Aarne-Thompson-Uther (ATU) index is the most widely used system in the field. The ATU index classifies tales into categories and names them with numbers based on motif. The *Cinderella* story, for example, is indexed and abstracted as follows:

> **510A *Cinderella***. (Cenerentola, Cendrillon, Aschenputtel.) A young woman is mistreated by her stepmother and stepsisters [S31, L55] and has to live in the ashes as a servant. When the sisters and the stepmother go to a ball (church), they give Cinderella an impossible task (e.g., sorting peas from ashes), which she accomplishes with the help of birds [B450]. She obtains beautiful clothing from a supernatural being [D1050.1, N815] or a tree that grows on the grave of her deceased mother [D815.1, D842.1, E323.2] and goes unknown to the ball. A prince falls in love with her [N711.6, N711.4], but she has to leave the ball early [C761.3]. The same thing happens on the next evening, but on the third evening, she loses one of her shoes [R221, F823.2].
>
> The prince will marry only the woman whom the shoe fits [H36.1]. The stepsisters cut pieces off their feet in order to make them fit into the shoe [K1911.3.3.1], but a bird calls attention to this deceit. Cinderella, who had first been hidden from the prince, tries on the shoe and it fits her. The prince marries her.
>
> (Christine and Koppy 2021, pp. 52-53)

The indexed item above shows *Cinderella* contains various elements of tales and represents a common element of the plot. The ATU index renders researchers to identify the elements of a certain plot and its commonalities from apparently variant tales.

Based on, or relating to, the ATU index, recent studies performed quantitative analyses of folktales, using various computational methods and statistical measurements (Bizzoni et al. 2022, Darányi et al. 1996, Nakawake et al. 2019, Roos and Heikkila 2009, Tehrani et al. 2016). To briefly survey the pioneering studies drawing on quantitative approaches and the ATU index, Tehrani et al. (2016), using 64 variants, analyzed the process of the formation of the tale type where ATU 123 (*The Wolf and Kids*) and ATU 333 (*Little Red Riding Hood*), for one example, were likely to have merged into the East Asian variants during the transmission. Darányi et al. (1996) examined the tales of magic, with 219 tale types and 1202 motifs, for multiple motif co-occurrences and identified the global pattern of most frequent motif sequences where they found the triplets and quadruplets being realized together in the folktales. Ofek et al. (2013) developed a tale type classification model based on bioinformatics mining methods to estimate and categorize the nature of tale types. Their model specified a set of motif sequences that may characterize the nature of magic tale types and further argued for the existence of consistent and stable motif sequences or what they call "narrative DNA." D'Huy (2013) shed light on the words that co-occur most frequently in proximity where, for one instance, the 'giant' emerged as a pivotal word and has its unique location in the lexical groups such as 'eye,' 'sheep,' 'man,' and 'cave,' and provided an intuitive access to the folkloric corpus by visualizing the word clusters.

One trend of recent quantitative research in humanities is to check quantitatively the validity of some well-known knowledge and to see if the analyses based on the quantitative methods bear out the observations of public understanding or previous academic studies. For instance, Nakawake et al. (2019) re-examined the intuitive folkloric knowledge of the predator-prey relationship (e.g., wolf-pig or cat-mouse relationship) based on natural language processing (NLP) of the 382 animal folktales. To measure the counts of animal occurrence, they extracted the nouns classified as 'animal,' and used WordNet, a lexical database of English and semantic components, and

confirmed that the paired animals, such as pig and wolf or mouse and cat, tend to co-occur, and that such adversarial relationship is likely to exist when deception motif is involved in the tales.

Da Silva et al. (2016) applied phylogenetic methods to folkloric corpora to measure the spatial, linguistic, and demographic distributions of folktales and to see how distant in time shared tales can be traced back in Indo-European oral traditions. Their research indicated that similar folkloric patterns are correlated with population histories and geographical proximity, thus supporting the long held theories concerning the relationships between folkloric traditions, geographical proximity, and the population histories. In the same vein, Acerbi et al. (2017) performed a statistical analysis on the relationship between folktale complexity and demography, using 380 animal tale types and 276 magic tale types, and measured the relationships between the number of the population, the number of tale types, the number of motifs, and the number of narrative traits. Their studies indicated that the population size of the region where the tales are collected is positively associated with the number of tale types and negatively associated with the number of motifs.

Nielbo et al. (2022) examined whether the automatic sentiment analysis of a fairytale correlates with its quality as perceived by human readers. They conducted the adaptive fractal analyses of 126 H.C. Andersen's fairytales and the ratings on GoodReads, a well-known online platform that grades and recommends books, and found a correlation between a tale's coherent pattern in sentiment and its quality perceived by human readers and further advocated for an extensive use of multifractal theory in the research of sentimental patterns in literary fields.

As is seen above, most of the previous quantitative analyses of folktales focus on specific sectors such as relationship, geographical characteristics, population, or readers' perception. Few studies, however, employed quantitative and qualitative methods in tandem with respect to the structure and elements of folktales. In this paper, we explore common character and event variables through statistical analyses with qualitative interpretation added, in order to find out the most important story elements in forming the respective type of LRRH. In addition, we discuss universal features of the structure of folktales in comparison with linguistic universals.

3. Method

This study involves one of the most widely known but debated tale types in folktale studies. The LRRH tales include ATU 333, ATU 123, and other unspecified types. Though conveniently subcategorized, it is difficult to draw absolute lines between ATU 333 (*Little Red Riding Hood*), ATU 123 (*The Wolf and the Kids*), and the unspecified type (for instance, Korean *The Sun and the Moon*). To briefly illustrate, shown are the examples of story elements, or what folklorists call motifs with an alphabet followed by a number as in N788. They are believed to have been the same story at a certain point of emergence or to be well grouped under the type umbrella by folklorists while showing a wide spectrum of details and differences along the story lines.

Table 1. Comparison of the LRRH Tale Types[2]

ATU 333 example		ATU 123 example		Unspecified example	Motif Index
Little Red Cap (Germany)	**Little Red Riding Hood (France)**	**The Wolf and the Seven Young Goats (Germany)**	**The Wolf and the Kid (Greece)**	**The Sun and the Moon (Korea)**	
A kid with food	A kid with food	Mother with food	Mother with food	Mother with food	N788
Wolf's disguise to grandmother	Wolf's disguise to grandmother	Wolf's disguise to mother	Wolf's disguise to mother	Tiger's disguise to mother	K2011
Hand test	Hand test	Voice, foot test	Voice, foot test	Voice, hand test	K1832/1839.1
Wolf's swallowing the victim	Wolf's swallowing the victim	Wolf's swallowing the victim	Wolf's failure	Tiger's getting into house	F911
Vomiting, Rescue from the Wolf	-	Vomiting, Rescue from the Wolf	-	Rescue from the Tiger	F913
Wolf's being drowned	-	Wolf's being drowned	-	The tiger's being fallen and dead	Q426

For this paper, we use the data collected and coded by Tehrani (2013), whose research on fairy tale analysis was a breakthrough in the field of folklore that used phylogenetic methods of measuring the lineages of tales. His indexing method is to parse a story into minimal and manageable narrative units to measure the distinction of the variant tales. His data included 58 variants of LRRH, and analyzed 72 story elements (aka motifs, or what he calls 72 variables), with character descriptors (number and sex of protagonist; type of the villain, wolf, ogre, and tiger) and event descriptors (building a house, going out of the house, devouring a victim). Since our research focused on well-known types of French *Little Red Riding Hood* (German *Little Red Cap*) and of Aesopian *The Wolf and the Kid* (respectively, ATU 333 and ATU 123), we put together the rest of the tales under the term "Unspecified." In addition, we eliminated the variables that were found "not applicable" in more than 30 versions of the tales as they are not common across the versions we investigated. Our modified data, therefore, has 58 variants of LRRH and 55 variables. The examples of story elements (with variable numbers)[3] can be showcased as follows:

Table 2. The Examples of the LRRH Elements

8	The sex of the villain: [0] male [1] female
9	The relationship of the villain to the victim: [0] stranger [1] father [2] aunt/uncle [3] friend
10	The relative: [0] absent [1] grandmother [2] father [3] aunt/uncle [4] mother [5] son [6] godfather
11	The setting: [0] absent [1] woods [2] mountains [3] cave
12	Guardian builds a safe home: [0] absent [1] present
13	The child goes out: [0] absent [1] present
14	Guardian goes out: [0] no [1] get food [2] visit relative [3] attend a feast [4] visit doctor

[2] This table is a modified version of Park and Kim's (2018) analysis.
[3] All the story elements are listed in Appendix I.

A square-bracketed number is allotted for statistical analyses in the logit model, i.e., logistic regression that is often adopted for classification and prediction by means of estimation of event probability. After coding the story elements as above, we used Python Scikit-Learn Library to implement logistic regression.

Before presenting the outcome of the logistic regression analysis, we run Feature Importance (Gini Impurity), which estimates the importance of each of all the relevant input features and, thus, shows us which feature has a larger effect than the rest of the others. In other words, a feature with a higher score is taken to have a larger effect on the model's prediction.

4. Results

As shown below, we identified the top ten Feature Importance scores to explain and to see which features may be most relevant to the distinction of the types. Selected are the coefficient values for the top ten features as being important to type prediction.

Table 3. Feature Importance Results

ATU 333		ATU 123		Unspecified	
Factor	Importance	Factor	Importance	Factor	Importance
13	0.798054	45	0.724509	1	1.298181
3	0.758416	66	0.589204	68	0.912987
24	0.616514	59	0.539169	55	0.883778
10	0.607492	30	0.389588	49	0.741630
16	0.554509	58	0.332242	40	0.707123
32	0.541613	57	0.314578	14	0.699552
5	0.467536	12	0.254057	25	0.690325
63	0.370107	44	0.224232	23	0.599281
1	0.298951	17	0.130104	72	0.545026
9	0.254212	9	0.069748	42	0.526291

Each factor is listed in scoring order in each tale type in Table 3. Note that the higher the score is, the larger the effect of a factor is in type differentiation. We will come back to the order above later in comparison with the results of the logistic regression analysis.

Next, we conducted logistic regression to identify and analyze the story elements in forming the tale types. Among the three tale types, ATU 123 merits our attention. As shown in Table 4, R-squared (0.962) nears 1 and Prob (F-statistic) is less than 0.05 (>1.88e-05) in *p*-value, and the tale type yields 14 story elements that are crucial in forming the tale type. ATU 333 and the Unspecified type do not yield any meaningful results in the logistic regression model. See Appendix II for the complete regression results of the three tale types.

Table 4. Regression Results of Tale Type ATU 123 (OLS Regression Results)

Dep. Variable:	Type	R-squared:	0.962
Model:	OLS	Adj. R-squared:	0.856
Method:	Least Squares	F-statistic:	9.054
Date:	Tue, 27 Apr 2021	Prob (F-statistic):	1.88e-05
Time:	12:56:40	Log-Likelihood:	55.716
No. Observations:	58	AIC:	-25.43
Df Residuals:	15	BIC:	63.17
Df Model:	42		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
1	-0.3546	0.151	-2.355	0.033*	-0.676	-0.034
4	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
6	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
7	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
8	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
11	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
19	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
27	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
29	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
34	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
37	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
38	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
43	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
47	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028

Omnibus:	2.160	Durbin-Watson:	2.311
Prob (Omnibus):	0.340	Jarque-Bera (JB):	1.365
Skew:	0.322	Prob (JB):	0.505
Kurtosis:	3.387	Cond. No.	1.57e+16

To recap the major results with the variable numbers, we can capture statistically meaningful factors as follows:

(1) P value < 0.05
ATU 123: [1, 4, 6, 7, 8, 11, 19, 27, 29, 34, 37, 38, 43, 47]

ATU 123 yields 14 factors meaningful to the formation of the tale type. To list the 14 factors, event descriptors are eight, and character and setting descriptors were six. Character and setting elements are the sex of the victim, the relationship of the guardian to the victim, the species of the villain, the sex of the villain, and the setting. The event elements are the encounter with the villain, the villain's disguising his voice, the password test, the youngest sibling's being tricked, the villain's sharing a bed with the victim, the villain's getting into the house where the victims hide, guardian's return home, and the victim's tricking the villain.

Now let us get back to the Feature Importance scoring order as recapitulated in (2):

(2) Feature importance
ATU 333: [13, 3, 24, 10, 16, 32, 5, 63, 1, 9]
ATU 123: [45, 66, 59, 30, 58, 57, 12, 44, 17, 9]
Unspecified: [1, 68, 55, 49, 40, 14, 25, 23, 72, 42]

The top ten feature importance scores show that the events (9 event variables versus 1 character variable for each type) are more relevant to the distinction of the type than the characters in the formation of ATU 123 and the unspecified type.[4] In the case of ATU 333, events and characters contributed evenly to the formation of the tale type (5 event variables versus 5 character variables). The comparison suggests that ATU 123 and the unspecified type are more driven by events than characters in forming feature importance. In other words, they are highly plot-driven tales rather than character-driven ones.

Particularly noticeable are the disparities in variables between feature importance and major factors based on the logistic regression analysis based on the same tale type. In case of ATU 123, no variable overlap can be found between OLS regression results (1, 4, 6, 7, 8, 11, 19, 27, 29, 34, 37, 38, 43, 47) and the feature importance results (45, 66, 59, 30, 58, 57, 12, 44, 17, 9). It suggests that story elements specified in the sample tales or descriptor specifics ('sheep' as opposed to 'goat' for instance) as realized characters and events may be differentiated from story elements potentially important in deciding the tale type—the point that requires further explanation in the following discussion.

5. Discussion

Literary approach can offer valuable means of cross-checking the results of quantitative analyses. Close text analysis is a hallmark of literary studies. Literary scholars examine not only words, sentences, and meaning, but also implication and significance. What is said or stated explicitly may be less important than what is implied and what is intended. For example, wedding is not literally meaningful in itself in folktales, but it is important as a function in the plot or construction of a tale, because the hero's wedding to a princess, thus inheritance of the kingdom, functions as a reward for the hero, or because a king's wedding to a queen, thus introducing a stepmother to the hero, functions as hardship and immanent conflict. In this vein, wedding has the significance as a part of the plot—the pivotal node in the storyline.

Even though there are no explicit overlaps in the story elements of feature importance and while details vary depending on the types, we see a pattern emerging in the functions assigned to respective story elements. When we line up the story elements from the perspective of the plot and see them as a type of function in the tales, we can observe the consistent sequence match as shown below.

[4] See Appendix III for the detailed list of the variables

◀ PLOT

	Mission	Departure	Villainy	Test	Escape	Challenge	Victory	◀ FUNCTION
ATU 333	16	13	24	32	63			
ATU 123	12	14		30	45	57 58 59	66	◀ STORY ELEMENTS
Unspec.		14	23	25	49	55 56	68	

ATU 333	13. Child goes out 16. The child/guardian takes food to the relative 24. The villain kills the relative 32. Victim(s) questions the villain about strange appearance 46. Advent of Helper such as huntsman, father or townsfolk 63. Others rescued from the villain's stomach
ATU 123	12. Guardian builds a safe home 14. Guardian goes out 17. The warning given to the victims 30. The hand test 44. The villain falls asleep after the feast 45. Victim rescued 57. Guardian gives the remains of the child to the villain to eat 58. Guardian accuses other animals 59. Guardian challenges the villain 66. Villain tricked into falling
Unspecified	14. guardian goes out 23. The villain kills the guardian 25. The villain's disguise 40. Victim hears villain crunching bones 42. The villain devours the victim 49. Children climb to safety 55. Guardian interrogates the other animals to track down the villain 56. Guardian invites the villain for tea 68. Monster is transformed into a lesser existence

Figure 1. Plot, Function, and Story Elements of the Three Types

In terms of plot, all three tale types share distinctive function sequences. ATU 333 shares four functions with ATU 123 and another four functions with the unspecified types. The functions of challenge and victory were not essential for the formation of ATU 333, and in fact, some variants such as French *Little Red Riding Hood* and Italian *Red Hat* have a tragic ending where the victim is swallowed by the wolf. Therefore, it can be assumed that the functions of challenge and victory were a later addition to the tale type.

Although the story details of ATU 123 and Unspecified type are diverse and different, their sequence of departure-test-escape-challenge-victory became a dominant and persisting constituent of the tale types. Particularly dominant is the function of challenge where the victim maneuvers a scheme and counteract villainy. Arguably, ATU 123 and Unspecified type might as well have evolved from the same tale types or one tale type spun off another tale type.

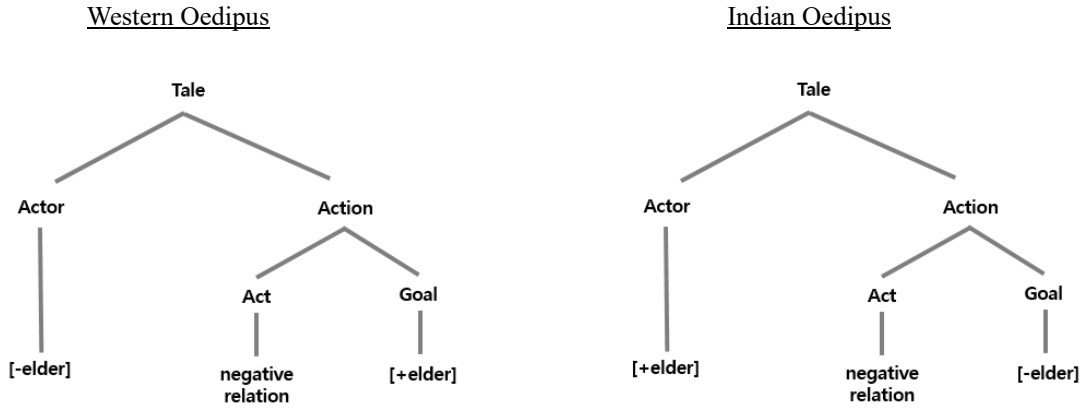
To extend our perspectives, we would like to take universality as one notable point from the current study. Propp

(1928) investigated one hundred Russian folktales and found that 31 elementary functions are reflected in actual events in the folktales. He argues that folktales are governed by common functions and similar plots. Shaul and Furbee (1998, p. 114) defined "plots as configurations of relations between characters" and showed a commonality of plots with the same relation between two characters but the different genders of characters in conflicts, based on the movie *Star Wars* and the fairy tale *Sleeping Beauty*;

Figure 2. Character Relations in *Star Wars* and *Sleeping Beauty*

In both stories, the protagonist is the son or daughter while the antagonist is the father or mother, and young ones overcome old ones. The nature of conflicting relationships and the ending of the tales carry universal features but show differences in the types of characters.

Referring to Claude Levi-Strauss' analysis of the Oedipus myth that Oedipus marries his mother whom he is closer to biologically and kills his father whom he is closer to psychologically, Ramanujan (1971) compares Western Oedipus type stories and Indian ones and found the following similarities and differences;



(cited from Shaul and Furbee (1998, p. 116))

Figure 3. Western Oedipus vs. Indian Oedipus

The universal features are found in terms of plot and relationship in the two different cultures, but mostly, protagonists are the younger in Western Oedipus while they are the elder in Indian Oedipus. We can easily identify universal features and parametric variation in tales across cultures, as we have done in the analysis of the LRRH type stories.

The universality issue we have delved into so far is very close to the analytical frames that linguists adopt when they try to figure out universal features in languages in the world. Let us take one of the principles proposed by generative linguists, which is X-bar theory. It accounts for phrase structures of syntactic categories of all the language and thus is a very powerful principle of the, so-called, Universal Grammar. It is illustrated in Figure 4,

where X is a variable that can be substituted with any syntactic categories such as N, V, A, Adv, P, and so on.

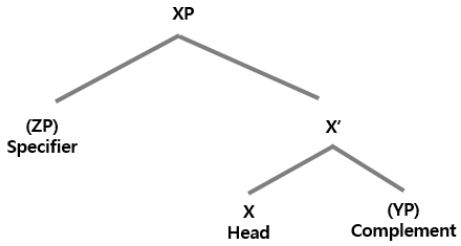


Figure 4. XP Structure

Generative linguists argue the configuration in Figure 4 can explain the structure of a phrase in every language, working as a universal principle, although there exists a parametric variation across languages with respect to the position of the head and its complement. For example, English is defined as a head-initial language in which the head precedes its complement while in Korean, a head-final language, the head follows its complement. X-bar theory renders us to treat phrase structures of both languages consistently, but variation resides in whether the head of a phrase precedes or follows its complement.

To wrap up, the current study of the LRRH type stories, the studies on the structure of folktales, and the linguistic study on phrase structures all commonly demonstrate there are universal properties and at the same time parametric variation in human language behavior. Placing the focus back on the quantitative analysis of the LRRH type stories, we emphasize the three types share common functions but bear different story elements.

6. Conclusion

In this article, we have shown that eight event descriptors and six character and setting descriptors are meaningful factors in the formation of ATU 123, on the basis of the statistical analysis of the three variant types of the LRRH stories, ATU 333. ATU 123, and Unspecified types. We have also presented all three types share certain functions even with variant forms of story elements. The universality observed in this study is on a par with what Maria Tatar asserts; despite variations and contradictions, “for all their rich variety, fairy tales have a remarkably stable—and therefore predictable—structure” and “beneath all the variations in its verbal realization the basic form still shines through” (Tatar 1987, xvi, xvii).

Investigating the three types of the LRRH tales, we found that each type lacks or has different forms of some story elements but their functions are common in all the types, obeying the same plot. We extended and linked our findings to the linguistic aspect to argue for universality in linguistic features and human language behavior. We expect future studies with other folktales will present more evidence of universality in the structure of folktales.

Finally, the results of this paper demonstrated that combined methods of quantitative and qualitative approaches can yield substantial insights into literary data which are difficult to solve by using statistical measures alone. Future studies can further illustrate that the combined measures can shed new light on the formation and reconstruction of folktales of the past.

References

Acerbi, A., J. Kendal and J. J. Tehrani. 2017. Cultural complexity and demography: The case of folktales. *Evolution and Human Behavior* 38(4), 474-480.

Bettelheim, B. 1976. *The Uses of Enchantment: The Meaning and Importance of Fairy Tales*. New York: Knopf.

Christine, K. and M. Koppy. 2021. *Fairy Tales in Contemporary American Culture: How We Hate to Love Them*. Lanham, MD: The Rowman & Littlefield Publishing Group, Inc.

Darányi, S., R. Zawiasa and Z. Hajnal. 1996. Conceptual mapping of a database in the humanities: First results of an experiment with Sophia. *Journal of Documentation* 52(1), 86-99.

Darnton, R. 1984. The meaning of Mother Goose. *The New York Review* (February 2, 1984), 41-47.

Da Silva, S. G. and J. J. Tehrani. 2016. Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *Royal Society Open Science* 3(1), 1-11.

d'Huy, J. 2013. A phylogenetic approach of mythology and its archaeological consequences. *Rock Art Research* 30(1), 91-103.

Nakawake, Y. and K. Sato. 2019. Systematic quantitative analyses reveal the folk-zoological knowledge embedded in folktales. *Palgrave Communications* 5(1), 1-10.

Nielbo, K., M. Thomsen, T. Peura and Y. Bizzoni. 2022. Fractal Sentiments and Fairy Tales: Fractal scaling of narrative arcs as predictor of the perceived quality of Andersen's fairy tales. *Journal of Data Mining & Digital Humanities.* https://doi.org/10.46298/jdmdh.9154

Ofek, N., S. Darányi and L. Rokach. 2013. Linking motif sequences with tale types by machine learning. In M. A. Finlayson, B. Fisseni, B. Löwe and J. C. Meister, eds., *Proceedings of the 4th Workshop on Computational Models of Narrative (CMN'13)*, Hamburg, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/OASIcs.CMN.2013.166

Park, J. and M. Kim. 2018. A study on development and application of motif-based storytelling. *Global Cultural Contents* 33, 23-45

Propp, V. 1928. *Morphology of the Folktale*. (Translated by L. Scott.) Austin, TX: University of Texas Press, 1968.

Ramanujan, A. K. 1983. The Indian Oedipus. In L. Edmunds and A. Dundes, eds., *Oedipus: A Folklore Case Book*, 234-264. New York: Garland Publishing Co.

Roos, T. and T. Heikkila. 2009. Evaluating methods for computer-assisted stemmatology using artificial benchmark data sets. *Literary and Linguistic Computing* 24, 417-433.

Shaul, D. L. and N. L. Furbee. 1998. *Language and Culture*. Prospects Heights, IL: Waveland Press, Inc.

Tatar, M. 1987. *The Hard Facts of the Grimm's Fairy Tale*s. Princeton, NJ: Princeton University Press.

Tehrani, J. J. 2013. The phylogeny of Little Red Riding Hood. *Plos One*. https://doi.org/ 10.1371/journal.pone. 0078871

Tehrani, J. J., Q. Nguyen and T. Roos. 2016. Oral fairy tale or literary fake? Investigating the origins of Little Red Riding Hood using phylogenetic network analysis. *Digital Scholarship in the Humanities* 31(3), 611-636. http://dx.doi.org/10.1093/llc/fqv016

Zipes, J. 1993. *The Trials and Tribulations of Little Red Riding Hood*, 2nd ed. New York: Routledge.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

Appendix I

The Story Elements of LRRH With Variable Numbers (Tehrani 2013)

1	Species of the victim: [0] animal [1] human
2	Type of animal: [0] goat [1] rabbit [2] duiker [3] sparrow
3	Victim is [0] multiple [1] single
4	Sex of the victim: [0] male and female [1] female [2] male
5	The victim wears a red cap/hood: [0] absent [1] present
6	Relationship of the guardian to the victim: [0] mother [1] brother [2] grandmother [3] father
7	The species of the villain: [0] fox [1] ogre [2] wolf [3] tiger/leopard [4] lion [5] bush beast [6] hyena [7] bear [8] alligator [9] crow
8	The sex of the villain: [0] male [1] female
9	The relationship of the villain to the victim: [0] stranger [1] father [2] aunt/uncle [3] friend
10	The relative: [0] absent [1] grandmother [2] father [3] aunt/uncle [4] mother [5] son [6] godfather
11	The setting: [0] absent [1] woods [2] mountains [3] cave
12	Guardian builds a safe home: [0] absent [1] present
13	The child goes out: [0] absent [1] present
14	Guardian goes out: [0] no [1] to get food [2] visit relative [3] attend a feast [4] visit doctor
15	The reason for visiting the relative: [0] not stated [1] borrow a skillet the mother wants to borrow [2] illness [3] birthday [4] select son-in-law [5] birth of a child
16	The child/guardian takes food to the relative: [0] absent [1] present
17	The instruction: [0] absent [1] children are warned not to stray from the path [2] not to open door [3] not to roast meat [4] bring back cakes
18	The child eats the contents of the basket: [0] absent [1] replaces them with donkey dung [2] nails
19	Encounter with the villain en route: [0] absent [1] the child encounters villain [2] the guardian
20	Reconnaissance: [0] absent [1] villain tricks victim into coming to its house [2] villain finds out where victim is going
21	The victim and villain take separate routes [0] absent [1] take the path of needles and pins [2] the villain takes the shortcut
22	The villain tricks the relative by posing as the child: [0] absent [1] present
23	The villain kills the guardian: [0] absent [1] present
24	The villain kills the relative: [0] absent [1] present
25	The villain's disguise: [0] absent [1] disguises as the guardian [2] disguises as the relative
26	The voice quality test 1 (villain disguises voice): [0] absent [1] present
27	Villain has an operation to clear his voice: [0] absent [1] present [2] rehearses the guardian's voice
28	Villain clears voice: [0] sitting on ant nest [1] eats/drinks something that changes voice [2] has tongue cut [3] hot coals to burn throat
29	The password rhyme test: [0] absent [1] present
30	The hand test: [0] absent [1] present
31	The villain disguises his paws: [0] absent [1] dye [2] banana leaves to make them smooth [3] wool threads [4] rolls in dirt [5] shells
32	Victim(s) questions the villain about strange appearance: [0] absent [1] present
33	Villain replies: [0] makes excuses about the journey [1] "all the better to eat you with!"
34	Youngest sib is tricked: [0] absent [1] youngest is tricked [2] youngest warns the others

35	Villain threatens the victim ”I’m coming closer!, closer!”: [0] absent [1] present
36	The villain's tail rattles in the basket: [0] absent [1] present
37	Monster goes to bed with the victim: [0] absent [1] present
38	The monster gets into the house and all the children hide in different rooms: [0] absent [1] present
39	The monster offers grandmother’s flesh to the victim: [0] absent [1] present
40	Victim hears villain crunching bones of the sibling: [0] absent [1] present
41	Victim does not understand the warning: [0] absent [1] the villain warns her [2] the cat warns her
42	The villain devours the victim :[0] absent [1] yes[2] puts victim in a sack
43	Guardian comes home: [0] absent [1] present
44	The villain falls asleep after the feast: [0] absent [1] present
45	Victim rescued: [0] by guardian [1] passerby
46	The hero: [0] absent [1] huntsman [2] father [3] townsfolk [4] peddler
47	Victim tricks the villain: [0] absent [1] letting her/them outside to urinate [2] to look at neighbor’s wedding
48	Villain ties rope/intestine round the victim’s foot: [0] absent [1] present
49	Children climb to safety: [0] absent [1] a tree [2] the roof
50	Villain sees the victim’s reflection in the pond: [0] absent [1] present
51	The children trick the monster into getting into a basket and drop him: [0] absent [1] present
52	Youngest sib tells monster to get an axe to climb up: [0] absent [1] present
53	Children grease the tree: [0] absent [1] present
54	Children assisted to escape the tree: [0] Gods let down a rope to heaven [1] passerby gives tokens [2] hangs clothes on the tree to fool the villain
55	Guardian interrogates the other animals to track down villain: [0] absent [1] present
56	Guardian invites the villain for tea: [0] absent [1] present
57	Guardian gives remains of child to the villain to eat: [0] absent [1] present
58	Guardian accuses other animals: [0] absent [1] present
59	Guardian challenges villain: [0] absent [1] fight [2]game
60	Guardian bribes the judge: [0] absent [1] present
61	The villain fights with fake horns: [0] absent [1] present
62	Rescued from the villain’s stomach: [0] absent [1] cut out of the monster's belly [2] freed from bag
63	Others rescued from the villain’s stomach: [0] absent [1] Guardian [2] other people inside monster are freed [3] relative freed
64	The villain is tricked into consuming dangerous substance: [0] absent [1] drinking hot oil [1] spear [2]
65	Monster stung by bees and wasps, scorpions, etc,: [0] absent [1] present
66	Villain tricked into falling: [0] absent [1] into the fiery pit [2] river
67	The monster's belly filled with stones: [0] absent [1] present
68	Monster is transformed: [0] absent [1] tree for honey bees [2] moon [3] cabbage
69	Children transformed into stars: [0] absent [1] present
70	The villain is killed by other monsters: [0] absent [1] present
71	Victim flees through the woods, and uses the help of the river, mountain, etc. to obstruct the villain’s pursuit: [0] absent [1] present
72	Victim hides until the morning, awaiting the villain's return: [0] absent [1] present

Appendix II

Logistic Regression Analysis of ATU 123, ATU 333, and Unspecified Type

ATU 123 (OLS Regression Results)

Dep. Variable:	Type	R-squared:	0.962
Model:	OLS	Adj. R-squared:	0.856
Method:	Least Squares	F-statistic:	9.054
Date:	Tue, 27 Apr 2021	Prob (F-statistic):	1.88e-05
Time:	12:56:40	Log-Likelihood:	55.716
No. Observations:	58	AIC:	-25.43
Df Residuals:	15	BIC:	63.17
Df Model:	42		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
1	-0.3546	0.151	-2.355	0.033*	-0.676	-0.034
3	-0.1480	0.162	-0.913	0.376	-0.493	0.198
4	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
5	0.1323	0.402	0.329	0.747	-0.725	0.990
6	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
7	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
8	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
9	-0.0881	0.141	-0.625	0.541	-0.388	0.212
10	-0.0706	0.048	-1.466	0.163	-0.173	0.032
11	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
12	0.3546	0.288	1.233	0.236	-0.258	0.968
13	-0.2347	0.250	-0.938	0.363	-0.768	0.298
14	0.0074	0.083	0.088	0.931	-0.170	0.185
16	0.0074	0.175	0.042	0.967	-0.365	0.380
17	0.0167	0.074	0.225	0.825	-0.142	0.175
19	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
23	-0.6610	0.419	-1.578	0.135	-1.554	0.232
24	-0.0586	0.343	-0.171	0.867	-0.789	0.672
25	-0.0724	0.123	-0.590	0.564	-0.334	0.189
26	0.0480	0.110	0.436	0.669	-0.186	0.282
27	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
29	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
30	0.0522	0.136	0.384	0.706	-0.237	0.342
32	-0.2181	0.182	-1.198	0.250	-0.606	0.170
34	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
35	0.0540	0.190	0.284	0.780	-0.351	0.459
36	-0.2884	0.272	-1.062	0.305	-0.867	0.290

37	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
38	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
39	0.2735	1.156	0.237	0.816	-2.191	2.738
40	0.3492	0.264	1.322	0.206	-0.214	0.912
41	-0.1979	0.500	-0.396	0.698	-1.265	0.869
42	0.0082	0.093	0.089	0.930	-0.189	0.206
43	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
44	0.0492	0.107	0.461	0.651	-0.178	0.277
45	0.0041	0.082	0.050	0.961	-0.170	0.179
46	0.0290	0.118	0.246	0.809	-0.222	0.280
47	-0.0793	0.024	-3.285	0.005*	-0.131	-0.028
49	0.1255	0.212	0.591	0.563	-0.327	0.578
50	0.2954	0.264	1.121	0.280	-0.267	0.857
55	-0.3095	0.182	-1.703	0.109	-0.697	0.078
56	-0.2139	0.151	-1.416	0.177	-0.536	0.108
57	0.2001	0.296	0.676	0.509	-0.430	0.830
58	0.0287	0.183	0.157	0.877	-0.361	0.418
59	0.0016	0.127	0.013	0.990	-0.268	0.272
60	-0.2270	0.257	-0.885	0.390	-0.774	0.320
62	0.0082	0.130	0.063	0.951	-0.270	0.286
63	-0.0302	0.066	-0.459	0.653	-0.170	0.110
64	-0.3385	0.248	-1.366	0.192	-0.867	0.190
65	-0.3878	0.292	-1.327	0.204	-1.011	0.235
66	0.0948	0.161	0.590	0.564	-0.248	0.437
68	-0.1657	0.106	-1.569	0.137	-0.391	0.059
70	-0.3850	0.407	-0.946	0.359	-1.252	0.482
71	-0.2422	0.637	-0.380	0.709	-1.601	1.116
72	-0.2226	0.334	-0.666	0.516	-0.935	0.490

Omnibus:	2.160	Durbin-Watson:	2.311
Prob (Omnibus):	0.340	Jarque-Bera (JB):	1.365
Skew:	0.322	Prob (JB):	0.505
Kurtosis:	3.387	Cond. No.	1.57e+16

ATU 333 (OLS Regression Results)

Dep. Variable:	Type	R-squared:	0.989
Model:	OLS	Adj. R-squared:	0.960
Method:	Least Squares	F-statistic:	33.60
Date:	Tue, 27 Apr 2021	Prob (F-statistic):	2.04e-09
Time:	12:56:40	Log-Likelihood:	95.435
No. Observations:	58	AIC:	-104.9
Df Residuals:	15	BIC:	-16.27
Df Model:	42		

Covariance Type:		nonrobust				
	coef	std err	t	P>|t|	0.025	0.975
1	-0.0093	0.076	-0.122	0.905	0.171	0.153
3	0.1477	0.082	1.808	0.091	-0.026	0.322
4	-0.0050	0.012	-0.410	0.688	-0.031	0.021
5	0.0930	0.203	0.458	0.653	-0.339	0.525
6	-0.0050	0.012	-0.410	0.688	-0.031	0.021
7	-0.0050	0.012	-0.410	0.688	-0.031	0.021
8	-0.0050	0.012	-0.410	0.688	-0.031	0.021
9	-0.0655	0.071	-0.922	0.371	-0.217	0.086
10	-0.0020	0.024	-0.082	0.936	-0.054	0.050
11	-0.0050	0.012	-0.410	0.688	-0.031	0.021
12	0.0866	0.145	0.597	0.559	-0.222	0.396
13	0.7274	0.126	5.768	0.000	0.459	0.996
14	-0.0784	0.042	-1.865	0.082	-0.168	0.011
16	0.0801	0.088	0.909	0.378	-0.108	0.268
17	-0.0050	0.037	-0.135	0.895	-0.085	0.075
19	-0.0050	0.012	-0.410	0.688	-0.031	0.021
23	0.1332	0.211	0.631	0.538	-0.317	0.583
24	-0.0748	0.173	-0.433	0.671	-0.443	0.293
25	-0.1192	0.062	-1.927	0.073	-0.251	0.013
26	0.0376	0.055	0.678	0.508	-0.081	0.156
27	-0.0050	0.012	-0.410	0.688	-0.031	0.021
29	-0.0050	0.012	-0.410	0.688	-0.031	0.021
30	-0.0226	0.068	-0.331	0.745	-0.169	0.123
32	0.0006	0.092	0.007	0.995	-0.195	0.196
34	-0.0050	0.012	-0.410	0.688	-0.031	0.021
35	-0.0029	0.096	-0.030	0.977	-0.207	0.202
36	0.2256	0.137	1.648	0.120	-0.066	0.517
37	-0.0050	0.012	-0.410	0.688	-0.031	0.021
38	-0.0050	0.012	-0.410	0.688	-0.031	0.021
39	0.2091	0.583	0.359	0.725	-1.034	1.452
40	0.0392	0.133	0.294	0.773	-0.245	0.323
41	-0.1522	0.252	-0.603	0.555	-0.690	0.386
42	0.0870	0.047	1.861	0.082	-0.013	0.187
43	-0.0050	0.012	-0.410	0.688	-0.031	0.021
44	0.0526	0.054	0.977	0.344	-0.062	0.167
45	0.0078	0.041	0.190	0.852	-0.080	0.096
46	0.0069	0.059	0.116	0.909	-0.120	0.133
47	-0.0050	0.012	-0.410	0.688	-0.031	0.021
49	0.1358	0.107	1.269	0.224	-0.092	0.364
50	-0.0014	0.133	-0.011	0.991	-0.285	0.282
55	-0.1034	0.092	-1.128	0.277	-0.299	0.092

56	-0.0186	0.076	-0.244	0.810	-0.181	0.144
57	-0.2147	0.149	-1.440	0.170	-0.533	0.103
58	0.0794	0.092	0.862	0.402	-0.117	0.276
59	-0.0097	0.064	-0.151	0.882	-0.146	0.126
60	-0.0996	0.129	-0.770	0.453	-0.375	0.176
62	0.0265	0.066	0.403	0.693	-0.114	0.167
63	0.0204	0.033	0.615	0.548	-0.050	0.091
64	-0.0958	0.125	-0.767	0.455	-0.362	0.170
65	-0.1781	0.147	-1.209	0.245	-0.492	0.136
66	0.0641	0.081	0.791	0.441	-0.109	0.237
68	-0.0971	0.053	-1.825	0.088	-0.211	0.016
70	-0.5953	0.205	-2.902	0.011	-1.032	-0.158
71	-0.1395	0.321	-0.434	0.670	-0.824	0.545
72	-0.2897	0.169	-1.718	0.106	-0.649	0.070

Omnibus:	3.814	Durbin-Watson:	2.031
Prob (Omnibus):	0.148	Jarque-Bera (JB):	2.839
Skew:	0.481	Prob (JB):	0.242
Kurtosis:	3.500	Cond. No.	1.57e+16

Unspecified type (OLS Regression Results)

Dep. Variable:	Type	R-squared:	0.948
Model:	OLS	Adj. R-squared:	0.802
Method:	Least Squares	F-statistic:	6.495
Date:	Tue, 27 Apr 2021	Prob (F-statistic):	0.000157
Time:	12:56:40	Log-Likelihood:	45.869
No. Observations:	58	AIC:	-5.738
Df Residuals:	15	BIC:	82.86
Df Model:	42		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
1	0.3639	0.178	2.039	0.059	-0.016	0.744
3	0.0002	0.192	0.001	0.999	-0.409	0.410
4	0.0074	0.029	0.258	0.800	-0.054	0.068
5	-0.2253	0.477	-0.473	0.643	-1.241	0.791
6	0.0074	0.029	0.258	0.800	-0.054	0.068
7	0.0074	0.029	0.258	0.800	-0.054	0.068
8	0.0074	0.029	0.258	0.800	-0.054	0.068
9	0.1536	0.167	0.920	0.372	-0.202	0.509
10	0.0726	0.057	1.272	0.223	-0.049	0.194
11	0.0074	0.029	0.258	0.800	-0.054	0.068
12	-0.4412	0.341	-1.295	0.215	-1.168	0.285
13	-0.4927	0.296	-1.662	0.117	-1.124	0.139

14	0.0711	0.099	0.719	0.483	-0.140	0.282
16	-0.0875	0.207	-0.423	0.679	-0.529	0.354
17	-0.0116	0.088	-0.132	0.897	-0.199	0.176
19	0.0074	0.029	0.258	0.800	-0.054	0.068
23	0.5277	0.496	1.063	0.305	-0.530	1.586
24	0.1334	0.406	0.329	0.747	-0.732	0.999
25	0.1915	0.145	1.317	0.207	-0.118	0.501
26	-0.0855	0.130	-0.657	0.521	-0.363	0.192
27	0.0074	0.029	0.258	0.800	-0.054	0.068
29	0.0074	0.029	0.258	0.800	-0.054	0.068
30	-0.0295	0.161	-0.183	0.857	-0.372	0.313
32	0.2174	0.216	1.008	0.330	-0.242	0.677
34	0.0074	0.029	0.258	0.800	-0.054	0.068
35	-0.0512	0.225	-0.227	0.824	-0.532	0.429
36	0.0627	0.322	0.195	0.848	-0.623	0.749
37	0.0074	0.029	0.258	0.800	-0.054	0.068
38	0.0074	0.029	0.258	0.800	-0.054	0.068
39	-0.4826	1.370	-0.352	0.730	-3.404	2.438
40	-0.3884	0.313	-1.241	0.234	-1.055	0.279
41	0.3501	0.593	0.590	0.564	-0.914	1.614
42	-0.0952	0.110	-0.867	0.400	-0.329	0.139
43	0.0074	0.029	0.258	0.800	-0.054	0.068
44	-0.1019	0.127	-0.805	0.434	-0.372	0.168
45	-0.0120	0.097	-0.123	0.904	-0.219	0.195
46	-0.0358	0.140	-0.257	0.801	-0.333	0.262
47	0.0074	0.029	0.258	0.800	-0.054	0.068
49	-0.2612	0.252	-1.038	0.316	-0.797	0.275
50	-0.2940	0.312	-0.941	0.362	-0.960	0.372
55	0.4129	0.215	1.917	0.075	-0.046	0.872
56	0.2326	0.179	1.299	0.214	-0.149	0.614
57	0.0146	0.350	0.042	0.967	-0.732	0.762
58	-0.1081	0.216	-0.499	0.625	-0.569	0.353
59	0.0081	0.150	0.054	0.958	-0.312	0.328
60	0.3267	0.304	1.074	0.300	-0.321	0.975
62	-0.0347	0.154	-0.224	0.826	-0.364	0.295
63	0.0098	0.078	0.126	0.902	-0.156	0.176
64	0.4344	0.294	1.479	0.160	-0.192	1.060
65	0.5659	0.346	1.634	0.123	-0.172	1.304
66	-0.1589	0.190	-0.834	0.417	-0.565	0.247
68	0.2628	0.125	2.101	0.053	-0.004	0.529
70	0.9802	0.482	2.033	0.060	-0.047	2.008
71	0.3817	0.755	0.505	0.621	-1.228	1.991
72	0.5123	0.396	1.293	0.216	-0.332	1.357

Omnibus:	0.074	Durbin-Watson:	2.378
Prob (Omnibus):	0.963	Jarque-Bera (JB):	0.207
Skew:	0.075	Prob (JB):	0.902
Kurtosis:	2.749	Cond. No.	1.57e+16

Appendix III

Event and Character Variables of Top Ten Feature Importance

ATU 333	Character variables	1. The species of the victim
		3. The number of the victim
		5. The victim wears a red cap/hood
		9. The relationship of the villain to the victim
		10. The identity of the relative
	Event variables	13. The child goes out
		16. The child/guardian takes food to the relative
		24. The villain kills the relative
		32. Victim(s) questions the villain about strange appearance
		63. Others rescued from the villain's stomach
ATU 123	Character variables	9. The relationship of the villain to the victim
	Event variables	12. Guardian builds a safe home
		17. The instruction
		30. The hand test
		44. The villain falls asleep after the feast
		45. Victim rescued
		57. Guardian gives remains of child to the villain to eat
		58. Guardian accuses other animals
		59. Guardian challenges villain
		66. Villain tricked into falling
Unspecified	Character variables	1. The species of the victim
	Event variables	14. Guardian goes out
		23. The villain kills the guardian
		25. The villain's disguise
		40. Victim hears villain crunching bones of the sibling
		42. The villain devours the victim
		49. Children climb to safety
		55. Guardian interrogates the other animals to track down villain
		68. Monster is transformed
		72. Victim hides until the morning, awaiting the villain's return