




ARTICLE

<https://doi.org/10.1057/s41599-019-0375-x>

OPEN

Systematic quantitative analyses reveal the folk-zoological knowledge embedded in folktales

Yo Nakawake^{1,2,4*} & Kosuke Sato ^{3,4*}

ABSTRACT Researchers have argued that folktales have the pedagogical function of transmitting adaptive information about the environment. Folk-zoological knowledge, such as the predator-prey relationship among wild animals or the relationships between wild and domesticated animals, is important knowledge for foraging and pastoral societies. Here, we analysed the descriptions of the 382 animal folktales listed in a worldwide tale type index (Aarne-Thompson-Uther type index) using natural language processing (NLP) and descriptive statistics. Our analyses suggested that first, the predator-prey relationship frequently appeared in a co-occurrent animal pair within a folktale (e.g., cat and mouse or wolf and pig), and second, the motif of ‘deception’, describing the antagonistic behaviour among animals, appeared relatively higher in ‘wild and domestic animals’ and ‘wild animals’ than in other types. Furthermore, the motif of ‘deception’ appeared more frequently in pairs, corresponding to the predator-prey relationship. These results corresponded with the hypothesis that the combination of animal characters and what happens in stories represent relationships in the real world. This study makes a new contribution by demonstrating that using a combination of quantitative methods and qualitative data to study folktales broadens our understanding of the evolutionary aspects of human cultures.

¹Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, UK. ²School of Human-Environmental Studies, Kyushu University, Kyushu, Japan. ³Organization for the Strategic Coordination of Research and Intellectual Properties, Meiji University, Meiji, Japan. ⁴These authors contributed equally: Yo Nakawake, Kosuke Sato. *email: yo.nakawake@anthro.ox.ac.uk; sato_kosuke@meiji.ac.jp

Introduction

Homo Fabulans. Folktales¹ are cultural universals, widely observed from hunter-gatherer societies to modern societies (Brown, 1991; Thompson, 1946), and popular stories such as ‘Cinderella’ or ‘Little Red Riding Hood’ have been found in cultural groups throughout the world (Tehrani, 2013; Zipes, 2012). A recent cultural phylogenetic analysis, a statistical method in evolutionary biology, revealed that worldwide folktales can be traced back 7000 years, even before the birth of writing systems (da Silva and Tehrani, 2016; Pagel, 2016). Traditionally, folktales were transmitted via verbal communications, but today folktales are transmitted and consumed via a wide range of media, including books, performing arts, and movies. Folktales are enjoyed as a form of entertainment in every type of society (Zipes, 1988, 2012). In hunter-gatherer societies, people spend considerable time around campfires at night, which provides an opportunity to share folktales (Dunbar, 2014; Wiessner, 2014), and anthropologists have argued that folktales contain various forms of knowledge essential for adaptation to both social and natural environments (e.g., Scalise Sugiyama, 2001; Smith et al., 2017).

One potential role of folktales is to increase sociality (Dunbar, 2014; Wiessner, 2014). While telling folktales is suggested to have the function of social grooming (Dunbar, 2014), empirical research has demonstrated that some folktales also contain social knowledge, such as moral lessons that can enhance cooperation (Smith et al., 2017). In the Agta community of the Philippines, folktales contain cooperative values (e.g., equality), which foster coordination and cooperation in foraging societies; furthermore, communities with skilled storytellers had higher rates of cooperation than other societies (Smith et al., 2017). These studies suggest that folktales have a pedagogical function to enforce social norms, which increases human adaptability to the local environment (Scalise Sugiyama, 2011, 2017).

The other potential role of folktales is to increase adaptability to the natural environment. Scalise Sugiyama has argued that many animal folktales contain folk-zoological knowledge about harmful predators and the nature of animals (Scalise Sugiyama, 1996, 2001, 2011, 2017). Although Scalise Sugiyama listed several cultural societies that transmit folk-zoological knowledge in folktales (Scalise Sugiyama, 1996), as far as we know, systematic and quantitative analyses have not yet been conducted on folk-zoological knowledge. In this study, we hypothesise that folk-zoological knowledge is embedded in folktales in such a way that the characters and events appearing in folktales represent the real-world environment. Below, we briefly explain how the transmission of folk-zoological knowledge can be adaptive.

Story and culture transmission. Knowledge about harmful animals is suggested to have an important role in human adaptation (Scalise Sugiyama, 1996, 2001), because harmful animals can directly attack humans or human property, including domestic animals. Some zoological knowledge, including the habitus of animals and predator-prey relationships (e.g., wolves and sheep) appears to be well-known and rather intuitive. However, these intuitive pieces of knowledge are extremely costly if they are learned individually via direct observation or through the experience of being attacked (Scalise Sugiyama, 1996, 2001). Listening to stories about, for example, an individual or an individual’s domestic animals being attacked can allow knowledge to be transmitted at a low cost (Scalise Sugiyama, 2001).² This ability to learn from others through stories is unique to humans (Boyd, 2018) because of the highly communicative skills that we possess, such as language or reading the intentions of others, which allow us to transmit information faithfully

(Tomasello, 1999; Boyd et al., 2011; Burdett et al., 2018). Empirical studies have demonstrated that cultural transmission through language allows knowledge to be transmitted faithfully through skill learning compared to other modalities (e.g., direct observation; Morgan et al., 2015). From these empirical observations in psychology or cognitive science, researchers have argued that storytelling and folktales have the function of transmitting adaptive knowledge effectively through language (Scalise Sugiyama, 2001, 2017; Zipes, 2012; Smith et al., 2017; Bietti et al., 2018; Boyd, 2018).

Another possible advantage of sharing folktales is cognitive attraction (Sperber and Hirschfeld, 2004; Scalise Sugiyama, 2017). Ethnographic evidence has demonstrated that folktales are often told in certain rhythms, with a ritualised style, or with redundancy (repetition of words) to attract children (Kroeber, 1948; Scalise Sugiyama, 2017; Tedlock, 1977; Wiessner, 2014). These ritualistic features of telling folktales or storytelling, in general, might have mnemonic advantages, making them easy to remember. Another form of cognitively attractive content is counterintuitive characters such as fairies or talkative animals, which are known to be highly memorable compared to intuitive characters (Barrett and Nyhof, 2001; Boyer, 2001). Counterintuitive characters are found in the majority of worldwide folktales (Barrett et al., 2009), and popular folktales such as the Brothers Grimm that tend to contain counterintuitive characters receive more attention on the Internet (Norenzayan et al., 2006). This also applies to modern stories such as urban myths (Stubbersfield and Tehrani, 2013). Zipes argued that some of the well-known folktales might have culturally evolved to become more cognitively attractive (Zipes, 2012); some folktales may acquire mnemonic features through transmission, guided by psychological bias or preference of content (Griffiths et al., 2008; Smith et al., 2008; see also Stubbersfield and Tehrani, 2013, who demonstrated that counterintuitive bias applies at the level of the whole story rather than at the concept level within a story).

Above, we discussed the functional or adaptive aspect of folktales; nevertheless, it is not always the case that beneficial cultural traits are spread through cultural transmission (Bentley et al., 2004; Lipo et al., 1997). For example, the distribution of cultural variants of the first name of newborn babies or archaeological pottery can be explained by value-neutral random copying (Bentley et al., 2004), which indicates that neither adaptive nor beneficial cultural traits frequently appear from random copying. In addition, in the case of conformist transmission (copying the cultural trait of the majority), neutral cultural traits may remain for several generations. If cultural variants influence the fitness of the traits being passed on, beneficial cultural traits are expected to evolve in the population; for example, hunting tools such as projectiles (Mesoudi and O’Brien, 2008a, 2008b) or hand axes (Kempe et al., 2012). However, folktales may be relatively less associated with immediate adaptive benefits compared to foraging technology. Thus, it is possible that the value-neutral cultural traits of folktales, such as being cognitively salient, are used for more than several generations, although this cognitive saliency may be beneficial to humans if the content of the folktales is adaptive. In this study, although the content of folktales might be value-neutral, we focus on the possible adaptive or functional aspects of folktales. Below, we briefly discuss how folktales have been studied in cultural evolution.

Cultural evolution of stories. While folktales can transmit cultural knowledge, the folktale itself may change its shape in the process of cultural evolution (Zipes, 2012; Boyd, 2018). When

researching cultural shifts, cultural evolution is used as an analogy of biological evolution by defining culture as the behavioural traits socially transmitted from one individual to another (Mesoudi, 2011). This approach has adapted the statistical methods of evolutionary biology to cultural studies. This means that using the analogy of biological evolution, various theories in the social sciences can be empirically tested, such as cultural shifts in the morphology of arrowheads (Mesoudi and O'Brien, 2008a) or religious rituals (Watts et al., 2016).

Folktales are a type of cultural material that has been studied rigorously in cultural evolution studies. The general theories of cultural evolution have been tested, such as how geographical and cultural boundaries influence cultural variations (Ross, Greenhill, and Atkinson, 2013) or how population diversity has influenced the variation of folktales (Acerbi et al., 2017). Furthermore, the application of phylogenetic methods on a certain type of folktale has made it possible to reveal its origin (da Silva and Tehrani, 2016; Pagel, 2016). This new quantitative approach opens the possibility of objectively investigating various questions about the source or period of the folktale.

Some cultural evolution research has used the textual content of folktales for analysis. For example, several studies analysed variations of 'Little Red Riding Hood' based on several features of the characters and events (e.g., whether the victim wears a red hood; Tehrani, 2013) or usage of articles (Karsdorp and Fonteyn, 2019), while other researchers examined the counter-intuitive content in urban myths and folktales (Norenzayan et al., 2006; Barrett et al., 2009; Stubbersfield and Tehrani, 2013). These studies have focused on either the content of one tale type or on one element (e.g., minimally counter-intuitive content) that was common across several folktales. However, relationships among the characters or elements have not been studied rigorously in a quantitative way to test the theory of cultural evolution.

In this study, we attempt to utilise semantic information with systematic and quantitative analyses to test the theory of cultural evolution. Specifically, we aim to reveal the knowledge embedded in folklore texts by analysing the Aarne-Thompson-Uther type index (ATU; Uther, 2004), a collection of international folktale abstracts that have been used in cultural evolution research (Ross et al., 2013; da Silva and Tehrani, 2016; Bortolini et al., 2017).

We investigated the folk-zoological knowledge embedded in folktales using co-occurrence and motif analyses. First, we conducted a co-occurrence analysis of animals. As we have argued that folktales contain folk-zoological knowledge useful to the real world, we hypothesised that the co-occurrence of the animals represents relationships in the real world. If a co-occurrent pair of animals in the story matched with the real-world relationship of animals, then the story was regarded as containing knowledge of the real world. However, the frequency of the co-occurrent pairs does not tell us the meaning of the pairs, and the interpretation was left to the researchers. Thus, second, we conducted a motif analysis to interpret the meaning of the animal pairs. As we hypothesised that the relationships of the animals represent adversarial relationships, in the motif analysis we focused on motifs that represent adversarial relationships. In order to investigate these hypotheses, we systematically analysed the descriptions of the animal folktales.

Methods

Corpus. We used the Aarne-Thompson-Uther type index (ATU) as our corpus, the set of text data to be analysed using natural language processing. The ATU is a catalogue of world folktales, and they are categorised by 'tale type', or the plot pattern of the tale. It was originally organised by the Finnish folklorist Antti Aarne, and further revisions were made by Stith Thompson and

Table 1 Index numbers and minor categories in animal tales (ATU 1-299).

Index	Minor category	N of tale types
1-99	Wild animals	110
100-149	Wild animals and domestic animals	72
150-199	Wild animals and humans	59
200-219	Domestic animals	35
220-299	Other animals and objects	106

Note that the category 'Other Animals and Objects' contains tales of animals not categorised as beasts (birds, bugs, or fish), inanimate objects, or imagistic characters. For details, see Supplemental Material (Table S1 for Top 30 animals in each category; Table S2 for inanimate objects)

then Hans-Jörg Uther (ATU is the abbreviation of the names of these editors). The feature of the ATU is that ATU numbers are assigned for each tale type. For example, the tale type 'fox and crane invite each other' was assigned ATU 60, and this is referred to as its index. In this research, we used Uther (2004), which is the latest English version of the ATU. We used the tale types labelled ATU 1-ATU 299, which are classified as 'Animal Tales'. These animal tales consist of five subcategories within the main category of animals (e.g., wild, domestic; see Table 1).

Although animals appear in other categories of folktales (e.g., 'Tales of Magic', labelled ATU 300-ATU 749), this study only focused on animal tales for the following two reasons. First, animal tales focus on the interaction between an animal and other animals or humans, and this is relevant to our main focus of the predator-prey relationship. Noy and Shenbar (2015) noted that in animal tales, the story takes place in the animal world, but they also noted that interactions between humans and animals appear in other categories where the story takes place in the human world. The second reason was a methodological reason. As stated above, animal tales have subcategories, such as 'wild' or 'wild and domestic', and as we were interested in utilising these subcategories, animal folktales provide the best tool.

The corpus contains 462 tale types. The number of indices and the total count of the tale types do not match, because some indices, such as ATU111, have multiple variants, such as ATU111a/ATU111b. Some indices are allocated to certain numbers but are yet to be allocated to certain folktale types. As we aimed to analyse the description of the folktales, we excluded tale types that had no description of tale content (e.g., they simply referred to another tale type such as 'see ATU XX'), thus 382 tale types with descriptions made up the final corpus to be analysed. To create machine-readable datasets of the corpus, we digitised a physical copy of the Uther (2004) edition using OCR (Optical Character Recognition). The scanned data were then checked and corrected manually.

Structure of the corpus. The structure of each tale type consists of (1) the title and content, (2) additional information, and (3) the bibliography and regional information.

(1) The title and content include the name of the tale type and a summary description of the story. This description is annotated with motifs for each time the texts of a description are matched to a certain motif. A motif is a repeated story element, such as a certain character, concept, or event that frequently occurs in the stories (Thompson, 1955). For example, motifs can be objects such as the devil or an angel, and they can be an action or event, such as the 'Creation of the Earth'. Each motif was tagged using the Thompson Motif Index (TMI), a systematic classification developed by Thompson (1955). TMI classified motifs into major categories labelled by letters from A to Z, and each subordinate

item was represented by a number (e.g., A1). For example, the letter ‘A’ stands for motifs related to ‘mythical motifs’ and TMI A2200–A2599 is the category group that represents the motifs that are ‘animal characteristics’, explanation of the origins of animals. While some tale types have more than one motif, there are some tales that have no motifs.

(2) Additional information includes the editors’ notes on whether the story (tale type) is part of famous collections such as ‘Aesop’s Fables’ and also whether the tale type is part of a combination with other tale types.

(3) Bibliography and regional information include where the folklore was recorded, as well as its bibliography.

Data processing tools. We extracted animal and motif information from the corpus. To extract data from the raw text file, we used Python (ver. 3.6.3) and the Natural Language Toolkit (NLTK; Bird et al., 2009) module (ver. 3.2.4). Information on the system requirements is available on our GitHub repository.

The occurrence of animals. To count animal occurrence, we processed the corpus in three steps. First, the animals were extracted from the corpus and listed. We extracted all of the words in the corpus and then extracted the nouns classified as ‘animal’. For classification, we used WordNet (ver.3.0) data included in the NLTK module. WordNet is a lexical database of the English language and records the semantic relationships of words (Miller, 1995). Using WordNet, we extracted the data of synsets from each word. A synset is group of synonyms, and it is tagged by lexical categories called lexicographers. There are 45 lexicographers, and they are expressed by combinations of the parts of speech tags, such as nouns or verbs, and categories of concept, such as person, act, or animal. We collected words that had the lexicographer ‘noun.animal’ in their synset(s), to extract the words related to animals. We made a list from these extracted words and verified it manually in the following way: if a listed word was not an animal’s name, then it was excluded. Words that can be signified as human (e.g., fisherperson), or imaginary characters remained on the list (e.g., dwarf).³

Second, according to the list we obtained, we counted the animals in the corpus for each tale type. The counted result was manually checked by RA, and miscounts and count omissions were corrected. In this procedure, verbs that could be confused with an animal name, such as ‘fly’, were removed from the count.

Finally, we unified the results of animals of one species that were labelled with different words (including ‘human’) and then coded these animal names. For example, cock, rooster, and hen were unified and labelled as ‘chicken’. Then, animal categories that appeared less frequently ($n < 5$) were unified (e.g., thrash, woodpecker, and parrot were labelled ‘bird’).

The co-occurrence of animals. Based on the list of animals counted using the method described above, we conducted co-occurrence analyses. To analyse the co-occurrence of the animals, we did not simply count the combinations and use the sum value, we also adjusted the value for the following reason. In one tale type, depending on the region, different animals are assigned to the same character role: for example, ‘fox’ in Europe can be substituted as ‘jackal’ in Africa (Uther, 2006). In the ATU, this substitution is indicated as brackets ‘fox (jackal)’ in the abstract. We should not interpret pairs, as a pair of ‘jackal’ and ‘fox’ appear together in the same folktale. Thus, we subtracted the number of all the substitutable pairs from the number of all of the combinations in the text.

Motifs. As described above, the Thompson Motif Index (TMI) classifies motifs into major categories using alphabetical letters (see Fig. 3 for each motif description), and we extracted TMI-tagged motifs in the corpus by using Python. In the analysis, we used the major categories represented by single letters (i.e., A to Z). Originally, the TMI classification system had a tree-like structure, and items in the same hierarchy were not necessarily in the same conceptual hierarchy (see Supplementary Table S7). Despite this hierarchy, we used the motifs (i.e., A to Z) as an indicator of a semantic relationship. We counted motifs such as ‘different word count’ in the NLP field. Even if the same motif appeared more than once in a tale type (e.g., K1 and K1), we counted it once but we counted more than once if the subordinate number was different (e.g., K1 and K2).

As we stated in the *Introduction*, we conducted the motif analysis to interpret the meaning of how the characters are represented in the folktale. From our hypothesis, if a pair of animals that appear in a folktale represent an adversarial relationship in the real world, then the folktale is expected to have a motif indicating this adversarial relationship. As such, we focused specifically on the motif ‘deception,’ which represents hostile relationships often appearing in nature as the predator-prey relationship, as predators and prey often deceive each other in the animal kingdom. We set the motif ‘deception’ as an indicator for general conflict, because ‘deception’ has various sub-categories that include ‘escape by deception (K 500–K699)’, ‘capture by deception (K700–K799)’, or ‘fatal deception (K 800–K999)’, and so represents general conflict in nature. Thus, we expected the motif ‘deception’ to appear more frequently in co-occurrent pairs in folktales about ‘wild animals’ and ‘wild and domestic animals’ but not in folktales about ‘domestic animals’.

Results

Co-occurrence analysis. To investigate the relationships between the animals, we calculated the animals’ co-occurrence frequency. Note that the co-occurrence frequency here refers to animals appearing at the same time in one tale type, and no duplicate count was made even if they appeared more than once in the tale type (see section: *The co-occurrence of animals in Methods*). The co-occurrence frequency of animals in the entire corpus is illustrated in Figs 1 and 2.

The figures demonstrate that common adversarial animal relationships, such as cat v. mouse, are often described as predator-prey relationships (Childs, 1986; Mahlaba et al., 2017). In addition, foxes were associated with poultry and small animals such as chicken or birds, and wolves were associated with large livestock such as pigs, sheep, and goats. These types of animal attacks have often been reported in livestock protection or food web studies (e.g., Ciucci and Boitani, 2012; Muhly et al., 2013; van Eeden et al., 2018). Thus, as we discussed in the introduction, pairs of wild animals and domestic animals appeared in the co-occurrence network, as represented in predator-prey relationships in the real world.

The frequency of occurrence of the motif. Above, we demonstrated that the pairs of animals, such as domestic v. wild, appeared in the co-occurrence analysis. Yet, this does not fully explain that the pairs of animal characters were in adversarial relationships. Thus, in order to infer the content of the story, we counted the frequency of the motifs (Fig. 3).

The most frequently appearing motif was K (‘deception’); this includes motifs where the character deceives another character, such as ‘deceptions through shams’ (K1700–2099) in ATU123 *The wolf and the kids*. The second most common motif was J (‘the wise and the foolish’), which includes motifs

when wisdom or stupidity leads to a consequence, such as the ‘absurd misunderstanding’ (J1750–J1849) in ATU34A ‘*The dog drops his meat for the reflection*’. The third most frequent motif was B (the animal motif), which includes motifs such as ‘animals with human traits’ (B200–B299) in ATU222 ‘*War between birds (insects) and quadrupeds*’. We considered the motif K (‘deception’) as reflective of the adversarial relationships of animals, such as in the animal kingdom when one animal deceives another animal (e.g., through mimicry or strategic deception).

The relationship between the classification of folk tales and motifs. In our hypothesis, we considered that folktales contain knowledge about the predator-prey relationship. Thus, the adversarial motif K implies that the predator-prey relationship should appear in accordance with real-world predator-prey relationships. First, we investigated this accordance of the motif and category of folktales, which contain information on whether

the animal characters were wild or domestic. In the corpus, the tale types were classified into five subcategories according to the type of animals that appear in the story: ‘wild animals’, ‘wild and domestic animals’, ‘wild animals and humans’, ‘domestic animals’, and ‘others’⁴. In order to investigate whether the distribution of the motifs was different in these subcategories, we calculated the relative frequency of the motifs in each subcategory and centred the values by subtracting the mean of each motif. Figure 4 illustrates that K (‘deception’) motifs appeared more frequently in ‘wild animals’ and ‘wild animals and domestic animals’ and less frequently in ‘domestic animals’ and ‘others’. In ‘domestic animals’, J (‘the wise and the foolish’) motifs appeared frequently, and in ‘wild animals and humans’, B (‘animal’) motifs appeared frequently.

Here, we focus on motif K, which we considered as a reflection of the predator-prey relationship. Motif K predominantly appeared in the categories ‘wild’ and ‘wild and domestic’, as most of the pairs of animals that appeared in the co-occurrence network were either among wild animals or wild v. domestic animals.

We further verified this point using a principal component analysis (PCA) to examine the variation pattern of the motifs according to the categories. We mapped the variation pattern in two dimensions in Fig. 5, and as our objective was to map these into fewer dimensions, we did not interpret the component of these axes. The cumulative contribution rate up to the second principal component was 0.92, and the variance of the data was almost explained by these two axes. In the figure, ‘wild animals’ and ‘wild animals and domestic animals’ were placed close to each other, and they were placed in the direction of the same

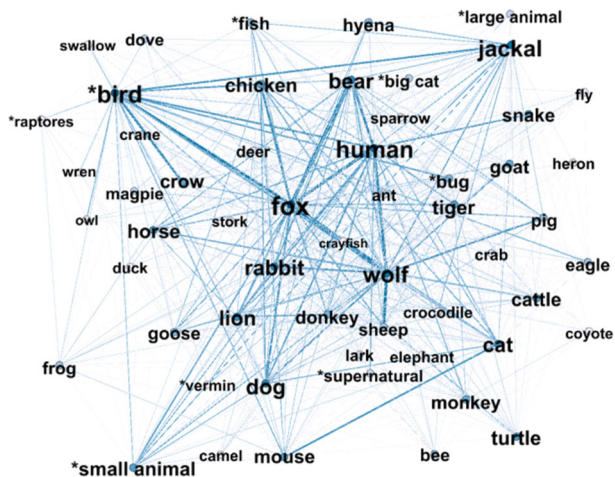


Fig. 1 Co-occurrence relationship of the animals. In this figure, the size of each label (e.g., dog or rabbit) represents the sum of the frequency of an animal that appeared co-concurrently. Each link between two labels represents a co-occurrent pair of animals, and the thickness of the links was adjusted by relative frequency. The labels with an asterisk (*) represent integrated categories of low-frequent animals. For example, “*bird” contains words such as ‘swan’ or ‘cuckoo’ that appear at a lower frequency, and it does not include ‘chicken’ or ‘crow’ appears frequently (see section: *The occurrence of animals in Methods*; and also see Supplementary Table S3 for the full list). Since Fig. 1 was informative, and so we then extracted the pairs of animals with a high frequency (Fig. 2. (a) $n > 10$; Fig. 2. (b) $n > 14$).

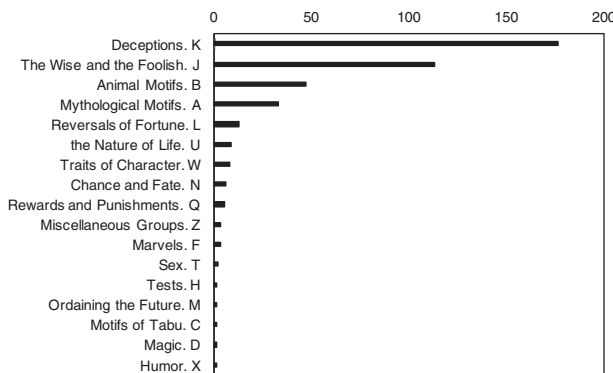


Fig. 3 Frequency of the motifs (Thompson motif index). We used the original label of motifs used by Thompson (1955).

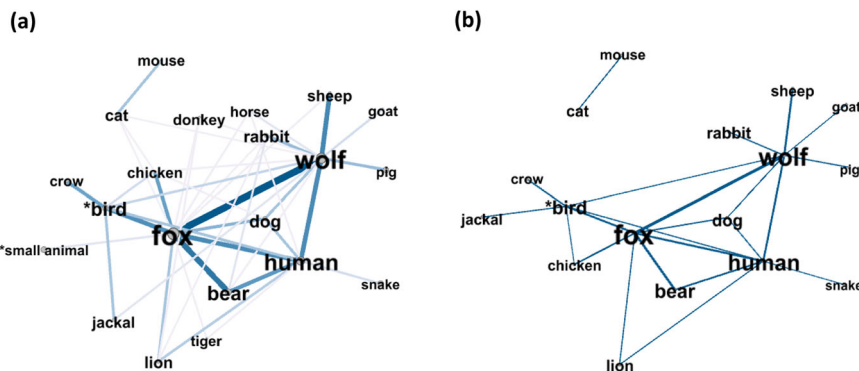


Fig. 2 Co-occurrence relationship of animals with a high frequency. **a** The co-occurrence network of animals with a frequency higher than 10. The density of the colours and width of the lines corresponds to the frequency of co-occurrence. **b** The co-occurrence network of animals with a frequency higher than 14.

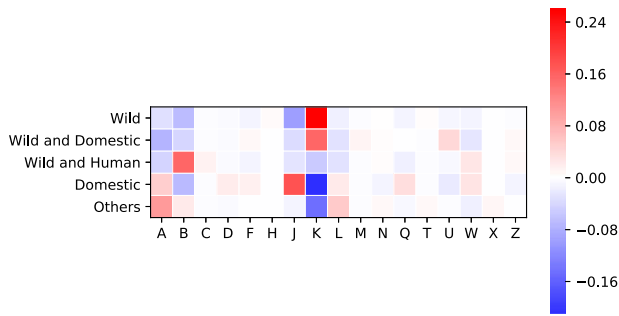


Fig. 4 Relative frequency of motifs by classification (after centring). The horizontal label A-Z represents the motifs (see Fig. 3 for details), and the vertical axis represents each category. The colour of each label illustrates the relative frequency of the motifs and the relative frequency within each classification were centred by subtracting the average values of each motif.

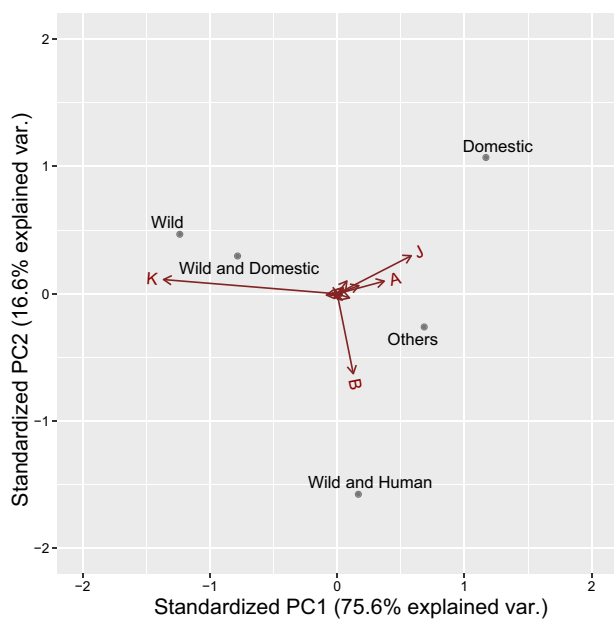


Fig. 5 Biplot of motif frequency by category.

vector K (‘deception’). Contrarily, ‘domestic animals’ was placed in the direction of vector J (‘the wise and the foolish’), and ‘wild animals and humans’ was placed in the direction of vector B (‘animal motifs’). The result of the PCA analysis was consistent with the relative frequency of the motifs according to the categories.

Relationship between individual animals and motifs. Our analysis detailed above was based on the categories in the ATU index. We further analysed the pattern of the individual animals to verify the consistency. We conducted a PCA on the animal occurrence with the motifs (Fig. 6). The cumulative contribution rate was 0.87 for the second principal component, and this explained the variance of the data sufficiently. Typical wild animals such as hyenas, jackals, foxes, and wolves, and their prey chickens, goats, and sheep were placed in a high position of the K (‘deception’) vector, which is consistent with the pairs in the co-occurrence network (see section *The frequency of occurrence of the motif in Results*). However, donkeys, horses, and cattle, three animals that demonstrated the direction of J (wisdom and foolishness) did not have a strong relationship in the co-occurrence network (they appeared in Fig. 2. (a) [$n > 10$] but not in Fig. 2. (b)

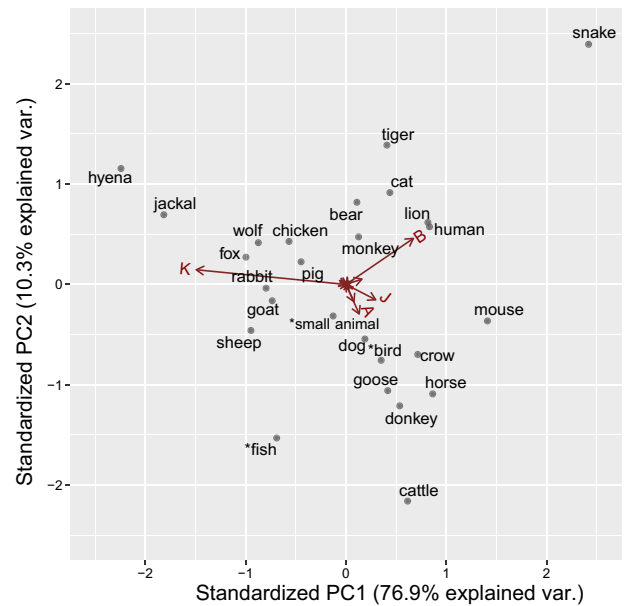


Fig. 6 Biplot of the frequency of the animals and motifs (animals where freq. >30).

[$n > 14$]). The series of analyses of the motifs and categories (relative frequency and PCA) demonstrated that a certain category was associated with specific motifs.

Further, the PCA on individual animals demonstrated animals that frequently appeared in pairs were associated with motif K. This suggested that adversarial pairs of animals frequently appeared in the folktales, similar to the natural environment (see also Supplementary Fig. S1).

Discussion

This study quantitatively analysed folktales in the ATU index by focusing on (1) the co-occurrence network of the animals, and (2) motif analysis of the categories and individual animals. The animal pairs that appeared in the co-occurrence network fit well with the real pattern of predator-prey relationships in the real world (e.g., fox v. chicken; wolf v. sheep). Furthermore, the motif analysis revealed that the relative frequency of the motifs was different across the folktale categories (i.e., domestic, wild, wild-domestic, wild-human, and others). This discrepancy of the relative frequency appeared especially in K (‘deception’) and J (‘the wise and foolish’). The motif K frequently appeared in the category of ‘wild’ and ‘wild and domestic’, and the motif J frequently appeared in the category of ‘domestic’. The result was consistent with the PCA of the motif by categories, and this further confirmed the results. Second, we conducted a PCA of the motifs at the level of individual animals, and this was consistent with previous categories. The results demonstrated that the predator-prey pairs of animals that appeared in the co-occurrence network (e.g., fox v. chicken, wolf v. sheep) was in the direction of the motif of adversarial relationships (K: ‘deception’). This suggests that the adversarial pairs of animals in predator-prey relationships frequently appeared in folktales.

Further discussion of motifs: difference between K and J.

Above, we considered that motif K (deception) was adversarial but not J (the wise and the foolish). The two motifs seem to be similar in the sense that one character was fooled or behaved foolishly; however, they are different in the relationality of the characters. While deception (motif K) always requires at least two characters, the deceiving and the deceived, foolish actions (motif

J) can be achieved alone. Here, we illustrate this difference using specific examples from folktales and explain how motif K represents an adversarial relationship, but motif J does not.

First, we illustrate the dyadic advisory relationship in K. For example, in ATU 127A *'The wolf induces the goat to come down from a cliff and devours it'*, a wild animal (the wolf) deceives a domestic animal (the goat). Further, in ATU 126 *'The sheep chases the wolf (where the sheep pretends to eat the wolf)'*, the relationship is reversed, and a domestic animal (the sheep) deceives a wild animal (the wolf). This type of adversarial predator-prey relationship is common in the wild, and domestic animals can form one part of it. However, the stories describing the relationships among domestic animals are not of the same kind. Instead of motif K, stories among domestic animals are likely tagged with motif J. For example, in ATU 211 *'The two donkeys and their loads'*, two donkeys, which are domestic animals, appear but they are not in an adversarial relationship. In the story, two donkeys are each carrying a load, and one donkey goes into some water. As it is carrying salt, the salt dissolves and its load becomes lighter. The other donkey witnesses the event and attempts to mimic the other donkey's action in order to lighten its load as well. However, this donkey is carrying flour and so its load becomes heavier instead. Thus, although two domestic animals appear in the story, there is no conflict as with predator-prey relationships.

As we demonstrated above, the relation of animals in J and K is different. In motif K, the wild and domestic animals are in an adversarial relationship, which is not the case for domestic animals in motif J. This is because domestic animals are rarely in adversary relationships in the real world, and folktales of domestic animals may represent this pattern. In contrast, wild animals have adversary relationships, and 'deception' can be observed in the animal kingdom among predator-prey relationships. While in reality, wolves do not use language to hunt goats, nor do sheep pretend to eat wolves, various deceptions can be observed in the animal kingdom, such as mimicry (Barber and Conner, 2007) and distractive display (e.g., distracting the attention of predators from younger individuals; Armstrong, 2008). This also appears in subcategories of 'deception (K)' such as 'escape by deception (K 500–K699)', 'capture by deception (K700–K799)' or 'fatal deception (K 800–K999)'.

Further discussion of motifs: other motifs. So far, we have not discussed other motifs, such as motif A ('mythical motif') or B ('animal motif'). The relative frequency of A and B were also different according to the categories. Motif A ('mythical motif') frequently appeared in the category of 'domestic' animals. Most of the subordinate motifs were items in A 2200–2259, which were about 'various causes of animal characteristics'. For example, ATU 200B *'Why dogs sniff at one another'* explained the cause of animal habits. Although these explanations are different from the evolutionary cause of certain animal behaviours, the origin of the actions was still interesting. While we could not explain why this motif appeared in domestic animals specifically, one possibility is that as we need to manage domestic animals in everyday life, retaining the knowledge of the behavioural characteristics of domestic animals may be more important than retaining knowledge about wild animals.

For motif B, unlike motif A, we found various subordinate motifs such as B200–299 'animals with human traits' (e.g., ATU235*: *The animals quarrel*), B500–599 'services of helpful animals' (e.g., ATU201D*: *Dogs bark at the thieves*), B700–799 'fanciful traits of animals' (e.g., ATU184: *Monkeys always copy man*), and B800–899 miscellaneous animal motifs. As the large category 'animal motif' was varied for the subordinate

motifs and certain subordinate categories did not consistently appear, we could not interpret the relative frequency of motif B.

Adaptive implication of knowledge transmission through stories. Above, we discussed how co-occurrent networks and motif analysis suggested that folktales contain folk-zoological knowledge. Here, we would like to discuss the evolutionary implications of this result. Researchers have argued that some folktales have an adaptive function to promote cooperative behaviours and spread the knowledge of the natural environment, which is beneficial for foraging or agriculture (Scalise Sugiyama, 2001, 2017; Zipes, 2012; Smith et al., 2017; Bietti et al., 2018; Boyd, 2018). The present research revealed the folk-zoological knowledge embedded in folktales, including knowledge essential for survival. However, why are folktales used to transmit this knowledge?

In modern society, we have a wide variety of information sources, such as books, TV programmes, or the Internet. However, when we did not have either printing systems or systematic education, animal folktales may have played the role of holding and sharing knowledge about animals. From an early age, children learn about animals, and although they rarely observe real animals (such as wolves or lions), children are aware that wolves are dangerous carnivores (e.g., from folk tales such as 'Little Red Riding Hood'); and this is because of the repetitive motif that appears in folktales (Sperber and Hirschfeld, 2004; Zipes, 1993, 2012).

The reason that folktales have been used to transmit information is not only because we did not have any other options, but also because folktales have beneficial features for transmitting knowledge. Storytelling is more than simply sharing information in a declarative format such as 'Animal X is dangerous' for two reasons (Scalise Sugiyama, 2017). First, folktales often contain ritualised styles, such as repetition or rhythm (Kroeber, 1948; Scalise Sugiyama, 2017; Tedlock, 1977; Wiessner, 2014). These repetitive or rhythmic styles are a feature of infant-directed speech, known as 'motherese' (Fernald, 1989), which has been demonstrated as a method for attracting the attention of infants and children aiding in their learning (Singh et al., 2002). Second, folktales often contain attractive cognitive motifs, such as counter-intuitive or humorous events or characters. Animal folktales often include a combination of these cognitively attractive elements to increase memorability, and across cultures, animals with counterintuitive motifs are found in higher frequencies (Barrett et al., 2009).

Limitations. Our study has several limitations. One of the limitations of the current research is that folktales do not mirror the real world perfectly. As discussed above, folktales contain counter-intuitive elements such as talkative animals. Furthermore, some animals have been symbolised or caricaturized through cultural transmission (Sperber and Hirschfeld, 2004). For example, the fox or jackal are often known as 'tricksters' with deceptive behaviours (Berezkin, 2014). As experimental studies have demonstrated, certain peculiar traits tend to be exaggerated during transmission (Bartlett, 1932; Barrett and Nyhof, 2001). For example, the deceptive behaviours of foxes may simply reflect real-world behaviours toward livestock or other prey. However, this association between 'fox' and 'deception' may be reinforced through repetitive transmission through generations, and so the fox becomes a symbol for deception without context, such as in 'crafty as a fox'. In this study, we have not analysed these culturally symbolised aspects of folktales, which other anthropologists and folklorists have discussed (e.g., Sperber and

Hirschfeld, 2004; Berezkin, 2014). Thus, further analyses in this direction are encouraged.

Another limitation is the geographical biases. In this study, we only analysed animal tales in the ATU. While the ATU is an international folktale catalogue, most of the stories in the ATU were collected in Europe (d'Huy et al., 2017). This is true of the animal tales that we analysed, as they were mostly from Northern Europe (for detail, see Supplementary Material S4). In order to generalise our results, it is necessary to other corpora obtained from outside Europe, such as a collection of animal stories from East Asia (Seki, 1950) or other regions.

In the analyses, we used TMI as an indicator of the relationships between animals. One potential problem is the concept of motifs. The motif is defined as the '*smallest element in a tale having a power to persist in tradition*' (Thompson, 1946, p. 415). However, this definition is criticised as being vague and unclear, and the problem of classification has been raised (Louwerse, 1997). Dundes (1997) also raised several problems with the TMI, such as the problem of 'Ghost Entries' (some indices have no corresponding references) or the 'Euro-centric bias'. One specific problem that may affect our results is 'Overlapping'. The motifs have several semantically similar motifs, although they are classified as different major categories. In our case, there were several motifs that could overlap with 'deception (K)', such as 'deceptive invitation to feast (J1577)'. However, we found that these overlapping motifs scarcely appeared in our corpus ($n = 7$), thus we consider the impact on our results of the overlapping problem is relatively low (See Supplementary Tables S4 and S5 for the list of overlapping motifs, and also S6 for the full list of motifs in our corpus).

Another problem related to motifs is the listing of the tag classification of the motifs. We used the motifs tagged by the ATU editors, but this may evoke the machine learning problem of precision-recall problems (Meder et al., 2016). As the editors of the ATU were experienced folklorists, the precision (also known as validity of the labelled class) of the motifs is expected to be high. However, the recall (also known as sensitivity; i.e., the completeness of all of the possible motifs) is unknown. In other words, we missed potential motifs because they were absent from the corpus. One possible solution is to use automatic tagging. Some researchers have proposed methods to classify motifs in folktales automatically using the NLP technique (Karsdorp and Van den Bosch, 2013; Meder et al., 2016), which would help to solve this problem.

Future perspective. In addition to the statistical analysis, NLP is useful in the study of cultural evolution. For example, Morin and Acerbi (2017) combined the theory of cultural evolution and the NLP method to understand the shifts in emotional expressions in the nineteenth century. Similarly, quantitative approaches using NLP or other statistical methods have been introduced as 'computational folkloristics' (Abello et al., 2012), and they form a bridge between qualitative data and quantitative methods. This study also illustrated the effectiveness of the quantitative approach to qualitative data.

In this study, we used relatively simple and already well-used analysis methods, such as co-occurrent network analysis or PCA, compared to other rigorous NLP techniques. The co-occurrent network of characters has already been used to develop a quantitative approach to stories, for example, to investigate the centrality of characters in *Hamlet* (Moretti, 2013), and thus the techniques that we used are not new. However, these techniques have not yet been used rigorously in the cultural evolution of folktales or stories as far as we know. In this study, we demonstrate that these simple techniques still provide meaningful

messages in the field of cultural evolution. We think that this study opens the door for the computational analysis of stories for researchers who have not yet used computational or statistical methods.

In our study, we focused solely on animal characters. However, this is only a small portion of the potentially interesting topics to be explored. For example, topics such as 'supernatural agents' or 'morality' frequently appear as elements in folktales (Zipse, 2012), and they have been investigated in the cultural evolution of stories (Norenzayan et al., 2006; Stubbersfield and Tehrani, 2013; Smith et al., 2017; Stubbersfield et al., 2019). Unlike animal names, extracting moral content may not be as straightforward, as moral content can appear without explicit words related to morality, such as 'fairness' or 'justice'. In addition, whether the story (or character) is moral or immoral may depend on the intention of characters (Strawson, 1962; Cushman et al., 2013) or the ending of the story (e.g., a villain is punished or not punished; cf. Piaget, 1932). Identifying whether a character is good or bad may require NLP techniques, such as agency detection (Karsdorp et al., 2015), which is necessary to identify who is acting harmfully. Further, we also need anthropological or psychological theory to judge what kind of narrative may be judged as moral (Graham et al., 2013; Curry et al., 2018). Above, we used morality as one example of a possible topic that requires further techniques or possible collaboration from different research fields. We hope that this study will contribute to a broader acceptance of this methodology and offer a scaffold for further studies.

Recently, quantitative analyses of large data have advanced our understanding of the humanities. Historians and anthropologists have analysed the accumulation of archaeological and historical datasets to test theories in social evolutions (e.g., Turchin et al., 2018). Similarly, our research tested an anthropological theory using quantitative analyses on quantitative datasets that have been accumulated by folklorists. Rapid technology advances have enabled us to answer new questions by using a rich accumulation of knowledge in humanities and social science; thus, further work can incorporate interdisciplinary studies to answer potentially large-scale questions about cultural evolution.

Data availability

All the programming codes and the datasets we processed are available at GitHub <https://github.com/satocos135/animalfolktale-analysis> (The original text of Uther (2004) is not uploaded due to the copyright).

Received: 19 July 2019; Accepted: 28 November 2019;

Published online: 17 December 2019

Notes

- 1 The term folktale is often used as the English translation of the German term '*Märchen*' (also translated as 'fairy-tale'), a type of story set in a fantastic or unrealistic world (Thompson, 1946). The term folktale is also used '*loosely to cover the whole range of traditional oral narratives*' (Thompson, 1946, p. 14). Narrowly defined, a folktale is one type of story or narrative that can be classified as a fictional story occurring at any time or in any place, unlike legends or myths, which are considered factual stories that took place in the distant past and in a different or earlier world than today (Bascom, 1965). Although there are several definitions for 'folktale', here we use the term operationally to describe tales found in the 'folktale index' rather than restricting the definition or precisely pointing to certain types of stories.
- 2 The possibility of identifying predatory animals is likely to be limited, as identifying the species based on backward induction requires knowledge of the species. Individual learning can often be costly and requires at least the experience of being attacked or witnessing an attack. Social learning, on the other hand, does not require costly first-hand experience, as information can be acquired through observation or imitation (Boyd and Richerson, 1985; Henrich and Boyd, 1998).

- 3 These remained as they could still be considered agents, and it might be possible that some animals might be co-occurrent with these supernatural characters. As described in Supplementary Table S3, the frequency of these characters is less than five in each category of folktale, and this is not included in our results in Fig. 2 ($n > 10$), and Fig. 6 ($n > 30$).
- 4 In the ATU index, the category 'Other' contains animals such as birds or insects together with inanimate objects and humans (see Supplementary Tables S1 and S2). Thus, this category might influence the results in the category level. Therefore, we also conducted analysis of the individual animal level to investigate the consistency (see *Relationship between individual animals and motifs in Results*).

References

- Abello J, Broadwell P, Tangherlini TR (2012) Computational folkloristics. *Commun ACM* 55(7):60. <https://doi.org/10.1145/2209249.2209267>
- Acerbi A, Kendal J, Tehrani JJ (2017) Cultural complexity and demography: the case of folktales. *Evol Human Behav* 38(4):474–480
- Armstrong EA (2008) Diversionary display: part 2. the nature and origin of distraction display. *Ibis* 91(2):179–188
- Barber JR, Conner WE (2007) Acoustic mimicry in a predator-prey interaction. *Proc Natl Acad Sci* 104(22):9331–9334
- Barrett J, Burdett ER, Porter T (2009) Counterintuitiveness in folktales: finding the cognitive optimum. *J Cogn Cult* 9(3):271–287
- Barrett JL, Nyhof MA (2001) Spreading non-natural concepts: the role of intuitive conceptual structures in memory and transmission of cultural materials. *J Cogn Cult* 1(1):69–100
- Bartlett FC (1932) *Remembering: a study in experimental and social psychology*. Cambridge University Press, Cambridge
- Bascom W (1965) The forms of folklore: prose narratives. *J Am Folklore* 78(307):3
- Bentley RA, Hahn MW, Shennan SJ (2004) Random drift and culture change. *Proc R Soc B* 271(1547):1443–1450
- Berezkin Y (2014) Three tricksters: World distribution of zoomorphic protagonists in folklore tales. In: Baran A, Laineste L, Voolaid P. (eds) *Scala Naturae: Stschriof in Honour of Arvo Krikmann for his 75th birthday*. ELM Scholarly Press, Tartu, pp 347–356
- Bietti LM, Tilston O, Bangerter A (2018) Storytelling as adaptive collective sense-making. *Topics Cogn Sci* 11(4):1–23
- Bird S, Klein E, Loper E (2009) *Natural language processing with python*. O'Reilly Media, Sebastopol
- Bortolini E, Pagani L, Crema ER, Sarno S, Barbieri C, Boattini A, Sazzini M, da Silva SG, Martini G, Metspalu M, Pettener D, Luiselli D, Tehrani JJ (2017) Inferring patterns of folktale diffusion using genomic data. *Proc Natl Acad Sci USA* 114(34):9140–9145
- Boyd B (2018) The evolution of stories: from mimesis to language, from fact to fiction. *Wiley Interdiscip Rev Cogn Sci* 9(1):1–16
- Boyd R, Richerson PJ (1985) *Culture and the evolutionary process*. University of Chicago Press, Chicago
- Boyd R, Richerson PJ, Henrich J (2011) The cultural niche: why social learning is essential for human adaptation. *Proc Natl Acad Sci USA* 108 (Suppl):10918–10925
- Boyer P (2001) *Religion explained: the evolutionary origins of religious thought*. Basic Books
- Brown DE (1991) *Human universals*. Temple University Press
- Burdett ERR, Dean G, Ronfard S (2018) A diverse and flexible teaching toolkit facilitates the human capacity for cumulative culture. *Review of philosophy and psychology*. *Rev Phil Psychol* 9(4):807–818. <https://doi.org/10.1007/s13164-017-0345-4>
- Childs JE (1986) Size-dependent predation on rats (*Rattus norvegicus*) by House Cats (*Felis catus*) in an urban setting. *J Mammal* 67(1):196–199
- Ciucci P, Boitani L (2012) Wolf and dog depredation on livestock in central Italy. *Wildlife Soc Bull* 26(3):504–514
- Curry OS, Mullins DA, Whitehouse H (2018) Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Curr Anthropol* 60(1):47–69
- Cushman F, Sheketoff R, Wharton S, Carey S (2013) The development of intent-based moral judgment. *Cognition* 127(1):6–21
- Dunbar RIM (2014) How conversations around campfires came to be. *Proc Natl Acad Sci* 111(39):14013–14014
- Dundes A (1997) The motif-index and the tale type index: a critique. *J Folklore Res* 34:195–202
- van Eeden LM, Eklund A, Miller JRB, López-Bao JV, Chapron G, Cejtin M, Crowther M, Dickman C, Frank J, Krofel M, Macdonald D, McManus J, Meyer T, Middleton A, Newsome T, Ripple W, Ritchie E, Schmitz O, Stoner K, Tourani M, Treves A (2018) Carnivore conservation needs evidence-based livestock protection. *PLoS Biol* 16(9):1–8. <https://doi.org/10.1371/journal.pbio.2005577>
- Fernald A (1989) Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Dev* 60(6):1497–510
- Graham J, Haidt J, Koleva S, Motyl M, Iyer R, Wojcik S, Ditto P (2013) Moral foundations theory: the pragmatic validity of moral pluralism. *Adv Exp Soc Psychol* 47:55–130
- Griffiths TL, Kalish ML, Lewandowsky S (2008) Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Phil Trans R Soc B* 363(1509):3503–3514
- Henrich J, Boyd R (1998) The evolution of conformist transmission and the emergence of between-group differences. *Evol Human Behav* 19(4):215–241
- d'Huy J, Le Quellec J, Berezkin Y, Lajoie P, Uther H (2017) Studying folktale diffusion needs unbiased dataset. *Proc Natl Acad Sci* 114(41):E8555–E8555
- Karsdorp F, Fonteyn L (2019) Cultural entrenchment of folktales is encoded in language. *Pal Commun* 5(1):25
- Karsdorp F, Van den Bosch A (2013) Identifying motifs in folktales using topic models. In: *Proceedings of the 22 Annual Belgian-Dutch Conference on Machine Learning, Nijmegen*, pp 41–49
- Karsdorp F, Van der Meulen M, Meder T, Van den Bosch A (2015) Animacy detection in stories. In: *Proceedings of the 6th Workshop on Computational Models of Narrative (CMN 2015)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, pp 82–97
- Kempe M, Lycett S, Mesoudi A (2012) An experimental test of the accumulated copying error model of cultural mutation for acheulean handaxe size. *PLoS ONE* 7(11):e48333
- Kroeber AL (1948) *Seven Mohave myths*. University of California Press (Anthropological records), Berkeley
- Lipo CP, Madsen M, Dunnell R, Hunt T (1997) Population structure, cultural transmission, and frequency seriation. *J Anthropol Archaeol* 16(4):301–333
- Louwerse M (1997) Bits and pieces: toward an interactive classification of folktales. *J Folklore Res* 34(3):245–249
- Mahlaba TAM, Monadjem A, McCleery R, Belmain S (2017) Domestic cats and dogs create a landscape of fear for pest rodents around rural homesteads. *PLoS ONE* 12(2):1–9
- Meder T, Karsdorp F, Nguyen D, Theune M, Trieschnigg D, Muiser IEC (2016) Automatic enrichment and classification of folktales in the Dutch folktale database. *J Am Folklore* 129(511):78–96
- Mesoudi A (2011) *Cultural evolution: how darwinian theory can explain human culture and synthesize the social sciences*. University of Chicago Press, Chicago
- Mesoudi A, O'Brien M (2008a) The cultural transmission of Great Basin projectile-point technology I: an experimental simulation. *Am Antiq* 73(1):3–28
- Mesoudi A, O'Brien M (2008b) The cultural transmission of Great Basin projectile-point technology II: an agent-based computer simulation. *Am Antiq* 73(4):627–644
- Miller GA (1995) *WordNet: a lexical database for English*. *Commun ACM* 38(11):39–41
- Moretti F (2013) *Distant reading*. Verso, London
- Morgan TJH, Uomini NT, Rendell LE, Chouinard-Thuly L, Street SE, Lewis HM, Cross CP, Evans C, Kearney R, de la Toore I, Whiten A, Laland KN (2015) Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat Commun* 6(1):6029
- Morin O, Acerbi A (2017) Birth of the cool: a two-centuries decline in emotional expression in Anglophone fiction. *Cogn Emotion* 31(8):1663–1675
- Muhly TB, Hebblewhite M, Paton D, Pitt J, Boyce M, Musiani M (2013) Humans strengthen bottom-up effects and weaken trophic cascades in a terrestrial food web. *PLoS ONE* 8(5):e64311
- Norenzayan A, Atran S, Faulkner J, Schaller M (2006) Memory and mystery: the cultural selection of minimally counterintuitive narratives. *Cogn Sci* 30(3):531–553
- Noy D, Shenbar A (2015) Animal Folktales. In: Patai R, Bar-Itzhak H (eds) *Encyclopedia of Jewish Folklore and Traditions*. Routledge, New York, pp 34–37
- Pagel M (2016) Anthropology: the long lives of fairy tales. *Curr Biol* 26(7):R279–R281
- Piaget J (1932) *The moral development of the child*. Kegan Paul, London
- Ross, R M, Greenhill, S J and Atkinson, Q D (2013) Population structure and cultural geography of a folktale in Europe. *Proc Biol Sci* 280(1756):20123065
- Scalise Sugiyama M (1996) Lions and tigers and bears: predators as a Folklore universal. In: Friedrich H, Jannidis F, Klein U, Mellmann K, Metzger S, Willems M (eds) *Anthropology and social history: Heuristics in the study of literature*. Mentis, Paderborn, pp 319–331
- Scalise Sugiyama M (2001) Food, foragers, and folklore: the role of narrative in human subsistence. *Evol Human Behav* 22(4):221–240
- Scalise Sugiyama M (2011) The forager oral tradition and the evolution of prolonged juvenility. *Front Psychol* 2:1–19
- Scalise Sugiyama M (2017) Oral storytelling as evidence of pedagogy in forager societies. *Front Psycho* 8:1–11
- Seki K (1950) *Nihon Mukashibanashi Shuhsei*. Vol. 1. Dobutsu Mukashibanashi. Sumikawa Shoten, Tokyo
- da Silva SG, Tehrani JJ (2016) Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *R Soc Open Sci* 3(1):150645

- Singh L, Morgan JL, Best CT (2002) Infants' listening preferences: baby talk or happy talk? *Infancy* 3(3):365–394
- Smith D, Schlaepfer P, Major K, Dyble M, Page A, Thompson J, Chaudhary N, Salali GD, Mace R, Astete L, Ngales M, Vinicius L, Migliano AB (2017) Cooperation and the evolution of hunter-gatherer storytelling. *Nat Commun* 8(1):1853.
- Smith K, Kalish ML, Griffiths TL, Lewandowsky S (2008) Introduction. Cultural transmission and the evolution of human behaviour. *Phil Trans R Soc B* 363 (1509):3469–3476
- Sperber D, Hirschfeld LA (2004) The cognitive foundations of cultural stability and diversity. *Trends Cogn Sci* 8(1):40–46
- Strawson PF (1962) Freedom and resentment. *Proc Br Acad*, 48(1):1–25
- Stubbersfield JM, Dean LG, Sheikh S, Laland NK, Cross CP (2019) Social transmission favours the 'morally good' over the 'merely arousing'. *Pal Commun* 5 (1):1–11
- Stubbersfield J, Tehrani J (2013) Expect the unexpected? Testing for minimally counterintuitive (MCI) bias in the transmission of contemporary legends: a computational phylogenetic approach. *Social Sci Comput Rev* 31(1):90–102
- Tedlock D (1977) Toward an oral poetics. *New Literary Hist* 8(3):507
- Tehrani JJ (2013) The phylogeny of little red riding hood. *PLoS One* 8(11). <https://doi.org/10.1371/journal.pone.0078871>
- Thompson S (1946) *The Folktale*. The Dryden Press, New York
- Thompson S (1955) Motif-index of folk-literature: a classification of narrative elements in folktales, ballads, myths, fables, mediaeval romances, exempla, fabliaux, jest-books, and local legends. Indiana University Press, Bloomington, Indiana
- Tomasello M (1999) *The cultural origins of human cognition*. Harvard University Press, Cambridge
- Turchin P, Currie T, Whitehouse H, François P, Feeney K, Mullins D, Hoyer D, Collins C, Grohmann S, Savage P, Mendel-Gleason G, Turner E, Dupeyron A, Cioni E, Reddish J, Levine J, Jordan G, Brandl E, Williams A, Cesaretti R, Krueger M, Ceccarelli A, Figliulo-Rosswurm J, Tuan P, Peregrine, P, Marciniak A, Preiser-Kapeller J, Kradin N, Korotayev A, Palmisano A, Baker, D, Bidmead J, Bol P, Christian D, Cook C, Covey A, Feinman G, Júlíusson Á, Kristinsson A, Miksic J, Mostern, R, Petrie, C, Rudiak-Gould P, ter Harr B, Wallace V, Mair V, Xie L, Baines J, Bridges E, Manning J, Lockhart B, Bogaard A, Spencer C (2018) Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proc Nat Acad Sci* 115(2):E144–E151
- Uther HJ (2004) The types of international folktales: a classification and bibliography, based on the system of Antti Aarne and Stith Thompson. *Suomalainen Tiedeakatemia, Academia Scientiarum Fennica (FF communications)*, Helsinki
- Uther HJ (2006) The fox in world literature: reflections on a 'fictional animal'. *Asian Folklore Stud* 65(2):133–160
- Watts J, Sheehan O, Atkinson Q, Bulbulia J, Gray RD (2016) Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature* 532 (7598):228–231. <https://doi.org/10.1038/nature17159>
- Wiessner PW (2014) Embers of society: firelight talk among the Ju/'hoansi Bushmen. *Proc Natl Acad Sci* 111(39):14027–14035
- Zipes J (1988) *The Brothers Grimm: from enchanted forests to the modern world*. Routledge
- Zipes J (1993) *The trials and tribulations of little red riding hood*. Routledge, London
- Zipes J (2012) *The irresistible fairy tale: the cultural and social history of a genre*. Princeton University Press, Oxford

Acknowledgements

We thank Dr. Christopher Kavanagh for commenting on an earlier version of this article, Dr. Rohan Kapitany for checking texts in GitHub, Dr. Masanori Takezawa for helpful comments for revising paper and Dr. Emily Burdett. We would like to thank Editage (www.editage.com) for English language editing, and research assistants for checking the lists of animals. The research was funded by Toshiba International Foundation.

Author contributions

The initial research was designed and conducted by YN and KS, who programmed and analysed the data. YN and KS wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1057/s41599-019-0375-x>.

Correspondence and requests for materials should be addressed to Y.N. or K.S.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2019