# Das konzeptuelle Modell des Zaubermärchens

## Einführung

Angesichts der unten aufgeführten axiomatischen Thesen streben wir einen sicheren Zugang zur Formalisierung des Zaubermärchens an und beabsichtigen folglich, die vollständige Automatisierung der Methode der vergleichenden Märchenforschung zu ermöglichen.

Die Thesen sind wie folgt:

1. Der Inhalt des Zaubermärchens setzt sich aus einer Abfolge einzelner Handlungen (Aktionen) zusammen.
2. Jede Handlung im Märchen wird von einer handlungstragenden Figur oder mehreren solcher Figuren geleitet.
3. Jeder Auftritt einer solchen Figur oder mehrerer Figuren löst eine neue Handlung aus.
4. Das Motiv, als Konzept für das kleinste, untrennbare Textteil, entspricht einer Handlung, die im Verlauf eines solchen Auftritts oder einer solchen Szene dargestellt wird.
5. Die Markierung der objektiv erkennbaren Szene im Text entspricht der Markierung der Handlungsabschnitte und somit auch der Markierung der Motive.
6. Die Kategorisierung der Handlungsabschnitte sowie der dabei beteiligten Figuren ist identisch mit der Definition inhaltlicher Konstanten für das gesamte Genre.
7. Die Schematisierung der gesetzmäßigen und folgerichtigen Reziprozität zwischen diesen Konstanten entspricht der Modellierung einer universellen Komposition des Genres.
8. Nach Abschluss der Schritte fünf bis sieben können wir die Maschine beauftragen, Märchen zu annotieren, auszuwerten und uns über die Ähnlichkeiten und Unterschiede zwischen den verglichenen Märchentypen zu berichten.

Die wichtigste Erkenntnis dieser Arbeit besteht in der Feststellung eines bisher fehlenden objektiven Kriteriums für die Erfassung des Motivs als kleinstes erzählerisches Inhaltselement im Zaubermärchen (Punkte 1-4).[[1]](#footnote-1)

Die vorliegende Arbeit beschreibt das aufgrund dieser Erkenntnisse entwickelte System von Handlungskategorien und den Kategorien handlungstragender Figuren. Das oberste Ziel der Formalisierung des Forschungsobjekts besteht in der Schematisierung der universellen Komposition des Zaubermärchens.

Der nächste Abschnitt der Arbeit widmet sich der Bereitstellung des Markups als Mittel zur standardisierten Kodierung bzw. Dekodierung des Inhalts des Märchens. Im Anschluss präsentiert die Arbeit einige Module des künstlichen Assistenten für die semiautomatische Annotation des Zaubermärchens sowie Tools für die visualisierte Darstellung der gegenseitig verglichenen Daten.

Generell repräsentiert diese Arbeit eine digitale Weiterentwicklung der Methode der Vergleichenden Märchenforschung und somit des wissenschaftlichen Erbes der internationalen Märchenforschung. Dabei stehen die Errungenschaften der Finnischen Schule und die von V. J. Propp erschlossene Morphologie des Zaubermärchens im Vordergrund.

## Handlungstragende Figuren im Märchen

Die Identifizierung und Kategorisierung der im Zaubermärchen anzutreffenden handlungstragenden Figuren erfordert zunächst die Auslegung der im Märchen funktionierenden Weltordnung. Basierend auf den Erkenntnissen empirischer Forschung lässt sich die in uns bislang bekannte Weltordnung in Zaubermärchen als hierarchisches System folgender Akteure beschreiben:

**Oberhaupt**

Das Oberhaupt kann am häufigsten ein König oder ein Familienoberhaupt sein, selten tritt an dieser Stelle auch Gott auf.

**Nachfolger**

Als Nachfolger gilt das handlungsfähige Kind des Oberhauptes.

**Nachwuchs**

Als Nachwuchs gilt der Vertreter der dritten Generation, zunächst das Kind des Nachfolgers.

**Ebenbürtiger vom außerhalb des Clans bzw. der Familie**

Die Ebenbürtige ist eine Person mit denselben Rechten wie der Nachfolger, gehört jedoch nicht dem Clan oder der Familie des Nachfolgers an.

**Priester / Magier / Totem**

Dieses kann durch ein besonderes Zeichen, eine Fähigkeit oder Fertigkeit hervorgehoben werden, wie zum Beispiel bärtig, betagt oder hässlich. Mager kann außerdem der Geist eines verstorbenen Familienmitglieds sein. Am häufigsten besitzen Totems oder personifizierte Naturerscheinungen wie z. B. Wind, Frost oder Sonne magische Kräfte. Vertreter dieser Kategorie stehen außerhalb der Macht von anderen Ordnungssubjekten derselben Welt.

**Untertan** **/ Berufstätiger / Beamte**

Zu den typischen Untertanen gehören Vertreter unehrlicher Berufe wie Müller, Hirt, Türmer oder Portier, aber auch andere Berufstätige wie Goldschmidt (wobei der Schmied eher mit einem Priester verglichen wird), Wirt, Kutscher, Kleinhändler; ebenso wie Beamte, zum Beispiel Minister, Offizier, Berater, usw. Wir sollen Untertan / Berufstätiger / Beamte als Stand vom Nachfolger mit demselben beruflichen Attribut unterscheiden. Ein gutes Beispiel wäre die häufige Verwendung eines Soldaten sowohl als Helden als auch als Untertan im Märchen.

**Besitz** **/ Gut /** **Sache**

Dazu gehören materielles und/oder immaterielles Erbe, das als Objekt eines Begehrens, Hilfsmittel oder Zaubermittel fungieren kann. Manchmal kann es auch ein Helfer sein, wie zum Beispiel ein Ring, eine Lampe, usw. Jedes Mitglied dieser Ordnung kann durch **männliches** oder **weibliches** Wesen vertreten werden, das gilt nicht selbstverständlich für den Besitz, das Gut und die Sache, es sei denn, dass sie personifiziert sind.

Diese Ordnung gilt für die beide Welten des Märchens. Diese werden nach der Perspektive des Helden als **eigene** und **fremde** Welten bezeichnet.

Funktion und Status der handlungstragenden Figur im Märchen werden durch ihre Zugehörigkeit zu verschiedenen Welten (eigene vs. fremde) sowie durch die dort vorhandenen Stände bestimmt. Insgesamt lassen sich folgende drei Kategorien von handlungstragenden Figuren identifizieren:

A. Protagonisten der eigenen Welt

B. Protagonisten der fremden Welt

C. Protagonisten der neutralen Welt (Figuren, die in beiden Welten zuhause sind).

Das Fremde bedeutet dabei nicht zwangsläufig das Feindliche; dennoch spiegelt jeder Konflikt im Zaubermärchen die Auseinandersetzung zwischen den Vertretern dieser beiden Welten wider.

Jeder Protagonist kann sich gegenüber seiner Welt sowohl freundlich als auch feindlich verhalten. Bei Protagonisten ohne feste Zugehörigkeit zu einer der Welten ist entscheidend, ob sie auf der Seite der Vertreter der eigenen Welt oder auf der Seite der fremden Welt stehen. Somit ergeben sich zwei Antipoden für jeden Protagonisten: richtig (r) und falsch (f).

Im Folgenden versuchen wir, die beschriebene Weltordnung und ihre konkreten Vertreter im Märchen in Form einer Matrix zu repräsentieren bzw. zu erklären:

Oberhalb der -xox Achse befindet sich die eigene Welt (die Welt des Helden). Der untere Bereich der -xox Achse entspricht der fremden Welt (die Welt des Antagonisten).

Rechts von der -yoy Achse befindet sich das Feld für richtiges Handeln, was bedeutet:

Die Vertreter der eigenen Welt handeln zugunsten ihrer eigenen Welt.

Die Vertreter der fremden Welt handeln zugunsten ihrer eigenen Welt.

Die Vertreter der neutralen Welt handeln zugunsten der Welt des Helden.

Links von der -yoy Achse befindet sich das Feld für falsches Handeln, was bedeutet:

Die Vertreter der eigenen Welt handeln zugunsten der fremden Welt.

Die Vertreter der fremden Welt handeln zugunsten der eigenen Welt.

Die Vertreter der neutralen Welt handeln zugunsten der fremden Welt.

Die -yoy Achse zeigt die Entitäten, die in beiden Welten funktionierende Ordnung beteiligt sind.

Die -xox Achse entspricht der Liste der Antipoden der Vertreter der Weltordnungsmitglieder im Märchen. Die Namen dieser Figuren sind jeweils mit zwei Großbuchstaben abgekürzt. Die Kleinbuchstaben 'r' bzw. 'f' vor den Abkürzungen weisen auf die Attribute richtig bzw. falsch des entsprechenden Protagonisten hin.

Die Initialen in Großbuchstaben sind wie folgt aufzulösen:

HH: Herr des Helden

HD: Held

RE: Rückeroberungsobjekt

HF: Helfer

ST: Stifter

VB: Verbindende Person

ZM: Zaubermittel

BZ: Besitzer des Zielobjektes

ZO: Zielobjekt

HP: Herr des potenziellen Partners

PP: Potenzieller Partner

AN: Antagonist

Die markierten Zellen in den Feldern und auf der -xox Achse zeigen den Bezug zwischen den Ständen der im Märchen vorhandenen Weltordnung und den handlungstragenden Figuren sowie ihre Zugehörigkeit zu eigenen, fremden oder neutralen Welten.

|  |  |
| --- | --- |
|  | Vertreter der eigenen Welt |
|  | Vertreter der neutralen Welt |
|  | Vertreter der fremden Welt |

Im Folgenden betrachten wir diese Figuren anhand ihrer richtigen und falschen Pendants.

Anmerkung: Es gibt einige Vorbehalte für die gesamte Nomenklatur der handlungstragenden Figuren. Die Berücksichtigung dieser Konventionen ist essenziell notwendig, um sprachliche Unstimmigkeiten zu vermeiden. Handlungstragende Figuren (Personen, Wesen, Gegenstände) stehen überall im Singular und werden durch das grammatische Genus des Terminus im Gegensatz zum Genus und Numerus der Instanz bezeichnet. Zum Beispiel wird für die weibliche Aschenputtel der Terminus rHD = richtiger Held verwendet; die beiden älteren Brüder werden als fHD = falscher Held bezeichnet, d. h. im Singular und im Maskulin, usw

**HH**  – Der Herr des Helden ist das Oberhaupt in der Familie. Wenn es sich um eine königliche Familie handelt, ist der Herr der König. Dieser muss vom König unterschieden werden, der der Herr des potenziellen Partners ist und nicht mit dem König gleichzusetzen ist, dem der Held dient.

Verhält sich der Herr des Helden unverhältnismäßig gegenüber dem Held, dann wird er als falscher Herr des Helden (fHH) definiert. Beispiele hierfür sind: Ein Vater begehrt nach seiner eigenen Tochter; er liefert seinen eigenen Sohn dem Verderben aus; oder er lässt seine Kinder in der Wildnis zurück. In einem anderen Szenario beauftragt ein König den jüngeren Sohn aufgrund des bösen Rats seiner älteren Söhne mit einer lebensgefährlichen Aufgabe. In all diesen Fällen übernimmt der Familienvater die Rolle eines falschen Herrn des Helden (fHH) und gleicht somit dem Antagonisten (s. unten).

**HD** – Der Held gilt als Maßstab für die Identifikation aller anderen Figuren im Märchen. In der Regel wird der Held als Nachfolger betrachtet, es sei denn, im Märchen gibt es keinen Herrn des Helden (HH). In umgekehrtem Fall, wenn im Märchen ein Protagonist dem Oberhaupt gleichgestellt ist, ist es notwendig, den Helden unter den anderen Protagonisten zu suchen. [[2]](#footnote-2) Des Weiteren kann es sowohl einen suchenden als auch einen vertriebenen Helden geben. Selten agieren im Märchen zwei Helden, die Geschwister sein können (auch Halbbruder oder Halbschwester) oder ein Nachfolger und ein Ebenbürtiger.

Als falsche Helden (fHD) gelten üblicherweise die Geschwister des Helden, sofern sie nicht vor dem Beginn der Heldentätigkeit geboren wurden. [[3]](#footnote-3) Die Rolle des fHD kann auch vom Ebenbürtigen übernommen werden. Dieser beginnt gewöhnlich mit der Schließung der Brüderschaft und endet mit dem Verrat gegenüber dem Helden.

**PP** – Der potenzielle Partner repräsentiert denselben Stand wie der Held, stammt jedoch aus der gegenüberstehenden Welt. Gewöhnlich hat der Partner einen eigenen Herrn (HP). Die Beziehung zwischen dem potenziellen Partner (PP) und dem Held (HD) beginnt immer aufgrund der Initiative einer der beiden Figuren, meistens durch die männliche Figur, selbst wenn die männliche Figur ein potenzieller Partner ist und nicht der Held, wie zum Beispiel im ATU 425. Daher ist es schwierig, eine weibliche potenzielle Partnerin von einem weiblichen Zielobjekt (ZO, siehe unten) zu unterscheiden. Letztere kann den Helden zwar heiraten, gilt jedoch nicht als potenzielle Partnerin für den Helden, sondern eher als Beute, die üblicherweise für den Auftraggeber, den falschen Herrn des Helden (fHH) oder den richtigen Antagonisten (rAN) bestimmt ist. Die Initiative, eine solche Figur zu gewinnen, geht vom rAN oder vom fHH aus.

Es ist ebenfalls schwierig, einen potenziellen Partner aus dem Rückeroberungsobjekt (RE) zu unterscheiden. Ein Identifikationsmerkmal ist die Heiratstüchtigkeit des Objekts, es muss also nicht der Familie des Helden angehören. Zum Beispiel gilt eine befreite Frau im ATU 301 als potenzielle Partnerin, auch wenn sie zuvor geraubt wurde und der Held den Auftrag oder die Erkundigung des Königs befolgt, um die entführte Frau zurückzuholen. Dieses Objekt wird weiterhin als potenzielle Partnerin betrachtet, selbst wenn der Held sie aus verschiedenen Gründen, einschließlich der oben genannten, nicht heiraten möchte.

Im Fall eines falschen potenziellen Partners (fPP) handelt es sich in der Regel um die gleichgeschlechtlichen Geschwister (manchmal auch Halbbruder oder Halbschwester) des richtigen potenziellen Partners oder seine Ebenbürtigen.

**HP** – Der Herr des potenziellen Partners repräsentiert denselben Stand wie der Herr des Helden, stammt jedoch aus der fremden Welt. Seine Hauptfunktion besteht darin, die Hochzeitsaufgaben und allgemein die Voraussetzungen zu stellen, und wenn nötig, die Freier zu verfolgen.

Der falsche Herr des potenziellen Partners handelt freundlich gegenüber dem Helden.

**RE** – Das Rückeroberungsobjekt (RE) kann sowohl eine Person als auch ein Gegenstand, ein Tier oder ein Zaubermittel sein. Es handelt sich um das Objekt, das vor dem Beginn der Aktivitäten des Helden in seiner Welt vorhanden war und nun zurückgeholt werden muss. Ist das Objekt nach dem Beginn der Aktivitäten des Helden weggenommen, verschollen oder verschwunden, behält es den Status bei, den es vor der Krise[[4]](#footnote-4) hatte, und muss nicht als RE betrachtet werden. (Beispielsweise bleibt der Schlange-Jüngling im ATU 425 auch dann potenzieller Partner, wenn er verschwunden ist und von der Gattin (HD) gesucht wird.) Das Vorhandensein des Rückeroberungsbedarfs bereits vor Beginn der Aktivitäten oder der Geburt des Helden oder des Helfers ist wichtig, um das gesuchte Objekt als Rückeroberungsobjekt (RE) zu definieren.

Das Rückeroberungsobjekt verhält sich in den meisten Fällen sehr aktiv. Der Held wird beauftragt, die entführte Mutter zurückzuholen. Sie hilft dem Helden dabei, den Aufenthaltsort der externen Seele des Antagonisten auszuspionieren.

Das Rückeroberungsobjekt gehört zur eigenen Welt und kann mit verschiedenen Ständen vertreten sein:

Nachfolger: Beispielsweise die Geschwister, die vom Antagonisten bereits vor der Geburt des Helden festgenommen wurden (ATU 328).

Oberhaupt: Wenn es sich beispielsweise um die Rückeroberung der entführten Mutter handelt, die jedoch ebenfalls vor dem Beginn der Aktivitäten des Helden entführt sein muss.

In einigen Fällen ist es schwer, das RE von der potenziellen Partnerin zu unterscheiden. Das Mädchen aus dem Typ ATU 301 kann sowohl Braut als auch geraubtes Objekt sein, das zurückgeholt werden muss. In solchen Fällen ist auf den Bezug zwischen dem betroffenen und anderen Protagonisten zu achten. Wenn ein Protagonist einen eigenen Herrn hat und sich in der fremden Welt befindet, handelt es sich um die potenzielle Partnerin.

Das RE kann dann falsch sein, wenn es sich gegen den Helden wendet. Wenn der Held seine von AN festgehaltenen Brüder befreit und diese ihn dann in den Brunnen werfen, sind die Brüder die Falschen Rückeroberungsobjekte (fRE) und nicht fHD (diese können sie sowieso nicht sein, wenn sie vor dem Beginn der Aktivitäten des Helden verschollen sind).

**ZO** u. **BZ**  – Das Zielobjekt und der Besitzer des Zielobjektes. Das Zielobjekt kann sowohl eine Person als auch ein Gegenstand, ein Tier oder ein Zaubermittel sein. Dieses Objekt gehört zu einer anderen Welt als der Held und wurde zuvor nicht erwähnt. Das Zielobjekt kann sich sowohl aktiv als auch passiv verhalten. Zum Beispiel der richtige Held (rHD) wird vom falschen Herrn des Helden (fHH) beauftragt einen Löwen (ZO) zu holen. Der Löwe leistet Widerstand, wird jedoch besiegt und mitgebracht. Der rHD wird vom richtigen Herrn des Helden (rHH) beauftragt, ein Pferd (ZO) zu holen. Der Held stiehlt das Pferd, das zwar inaktiv ist, aber dem rHD später andere Aufgaben zu lösen hilft, indem es als Helfer fungiert.

In den meisten Fällen haben diese Objekte ihre Besitzer. Der beauftragte Held kommt oft bei der Eroberung des Objekts unmittelbar mit dem Besitzer in Kontakt. Es ist üblich, dass der Besitzer selbst als Zielobjekt angefordert wird, was anschließend zur Bestrafung des Antagonisten führt (typische Beispiele ATU 328, 531). Die weibliche Figur des Zielobjekts ist von der potenziellen Partnerin (PP) auch dadurch zu unterscheiden, dass sie keiner Herrschaft untersteht. Nach ihrem Stand soll die weibliche Figur des Zielobjekts eine Magin bzw. ein Gegenstand oder Besitz des Magers sein. Dies bedeutet jedoch nicht, dass alle Figuren im Märchen, die mit magischen Eigenschaften ausgestattet sind, der Kategorie Zielobjekt angehören. Ein Beispiel hierfür ist das Märchen ATU 313, in dem der Jüngling (rHD) beim Zauberer (rHP) gerät und sich in dessen Tochter (rPP) verliebt, mit der er später vor dem rHP flieht. Ein Zielobjekt kann auch eine abstrakte Aufgabe sein, wie Acker pflügen, Holz holen, Zoll erheben usw., die nicht als Protagonisten des Märchens betrachtet werden.

Ein falsches Zielobjekt liegt vor, wenn das Objekt zugunsten des Helden handelt. Beispielsweise wird nach dem Organ eines Tieres gesucht, und das Tier schenkt dem Helden seinen Jungen. Die Schöne folgt dem Helden, ohne dabei Widerstand zu leisten, usw. Dass diese Objekte später dem Helden helfen, ist jedoch üblich. Im Fall von Zaubergegenständen könnte es sich um ein falsches Objekt handeln, wenn es ausgetauscht wird, zum Beispiel das Lebenswasser.

**ZM** - Ein Zaubermittel ist nur dann als solches zu definieren, wenn es nicht als Zielobjekt fungiert. Es muss also nicht gezielt gesucht und erworben werden, sondern kann geschenkt, bestellt, gefunden oder entwendet werden. Der Zauberring, den der Held als Belohnung für seine guten Taten erhält, ist kein Zielobjekt, sondern ein Zaubermittel (ZM). Im Gegensatz dazu ist das Lebenswasser, das der Held im Auftrag seines Vaters beschafft, das Zielobjekt (ZO) und nicht das Zaubermittel, obwohl es tatsächlich magische Eigenschaften hat. Wenn ein echter Ring durch einen falschen ausgetauscht wird, handelt es sich um das falsche Zaubermittel (fZM). Gleiches gilt für den Fall des ausgetauschten Lebenswassers (fZO).

**ST** u. **HF**  – Der Stifter und der Helfer. Der Stifter kann sowohl ein Mensch als auch Tier, eine Pflanzen oder mythisches Wesen sein. Der Stifter liefert dem Helden einen notwendigen Gegenstand, eine Fähigkeit oder Informationen, hat jedoch selbst keine direkte Wirkung auf den Helden. Beispielsweise kann der Stifter den Helden nicht beleben oder ihm Asyl bzw. Schutz gewähren. Dabei ist es wichtig, dass der Stifter dem Helden lediglich unterstützende und keinesfalls belastende Informationen liefert. Statt beispielsweise zu sagen, dass ein Ungeheuer eine Königstochter entführt hat, wäre es passender, dem Helden den Weg zum Entführer zu weisen oder ihm eine Waffe zu übergeben oder einen Trick zu zeigen, wie man das Ungeheuer besiegen kann.

Der Helfer ist jemand, der den Helden begleitet und gemeinsam mit ihm handelt. Das gemeinsame Wirken mit dem Helden bei der Lösung der Aufgabe oder Behebung der Notlage ist ein wichtiger Faktor, um den Helfer vom Stifter zu unterscheiden. Die Funktion des Stifters besteht darin, dem Helden zu helfen, aber so, dass er dabei nicht als Mitwirkender agiert. Manchmal können die Funktionen von Stifter und Helfer in einer Figur verflochten sein, wie zum Beispiel ein Tier-Schwager aus ATU 552. Er zeigt dem Helden zunächst den Weg zum Antagonisten und übernimmt somit die Funktion des Stifters. Dann jedoch, als der Held im Zweikampf mit dem Antagonisten getötet wird, erweckt ihn derselbe Schwager wieder zum Leben und erfüllt dadurch die Funktion des Helfers. Nach seinem Stand kann ein Helfer ein Mager oder ein Angehöriger des Magers sein.

Stiften und helfen können fast alle Protagonisten im Märchen, einschließlich des Antagonisten. Der Stifter oder Helfer selbst kann jedoch nichts anderes tun als zu stiften und zu helfen.

In der Welt der Mager und Totem gibt es keine Stände und Hierarchie: Eine junge Schlange wird gerettet, und der Vater der Schlange schenkt dem Retter einen Wunschstein. Eine andere Schlange wird gerettet, und sie beschenkt den Retter mit der Fähigkeit, die Tiersprache zu verstehen.

Der falsche Stifter (fST) bzw. falsche Helfer (fHF) ist jemand, der dem Bösen hilft oder etwas stiftet, zum Beispiel die von einem falschen Helden beauftragte Hexe. Auch die Wesen, die beim bösen Zielobjekt die Wache halten (ATU 551), gelten als falsche Helfer.

**VB**  – Die verbindende Person wird durch ein menschliches oder mythisches, jedoch eher anthropomorphes Wesen dargestellt. Ihre Hauptfunktion besteht darin, die Vermittlung zwischen dem Helden und einer anderen handlungstragenden Figur zu gewährleisten oder den Helden über die bestehende Krise bzw. notwendige Belastungen zu informieren. Typischerweise werden in der Rolle der verbindenden Person Untertanen eingesetzt, wie etwa eine alte Gastgeberin, ein Wirt oder ein Hirt. Oftmals ist auch der Herr des Helden, meistens die Mutter des Helden, in dieser Rolle anzutreffen. Sie kann zum Beispiel zum König gehen und ihm den Heiratsantrag ihres Sohnes überreichen. Ein weiteres Beispiel findet sich im Märchen ATU 425, wo die Schwester des potenziellen Partners dem Helden hilft und beim Händewaschen den Ring in die Hände der gesuchten Partnerin fallen lässt. In ATU 552 arrangiert die Schwester das Treffen mit dem Stifter.

Eine verbindende Person wird zur Falschen, wenn sie zu Gunsten des Antagonisten handelt. Häufig finden sich in dieser Rolle Diener, Berater, Portiers usw., die beim bösen König beschäftigt sind. Die Handlung der falschen verbindenden Person ist nicht immer das Gegenteil dessen, was die Richtige tut. Als typisch gilt das Denunzieren und die Rolle eines Denunzianten oder die Initiierung böswilliger Aufträge.

**AN**  – Antagonist (Gegenspieler, Schädling): Der Antagonist gehört zur bösen Welt (Minus-Welt) und kann praktisch durch alle Stände vertreten werden. Wie oben angedeutet, kann seine Funktion auch von einem anderen Protagonisten mit einem Minuszeichen, jedoch aus der Plus-Welt, übernommen werden. Besonders verbreitete Fälle sind der falsche Herr des Helden (fHH) und der falsche Held (fHD).

Wenn ein anderer Vertreter desselben Standes zu Gunsten des Helden handelt, kann es sich um einen falschen Antagonisten handeln. Ein Beispiel hierfür könnte die Frau des Bösen sein (dies ist jedoch eher eine Vermutung).

In einem Märchen können mehrere Antagonisten auftreten. Wenn der Drache getötet wird und seine Schwester versucht, Rache zu nehmen, ist auch die Schwester des Drachen eine Antagonistin.

## Handlungsabschnitte im Märchen

Die vorliegende Methode zur Segmentierung des Märchens basiert auf zwei objektiv erkennbaren Entitäten im Märchen:

* Die durch den gesamten Text präsentierte Vollgeschichte oder Story.
* Das durch den Auftritt einer handlungstragenden Figur erzeugte Motiv als kleinste Teilgeschichte in der Vollgeschichte.

Im Kontext der Märchensegmentierung ist die Bedeutung der zweiten Entität selbstverständlich und besonders wichtig. Wenn wir darauf achten, dass im volkstümlichen Märchen keine Porträts, Naturbilder oder Schilderungen von Gefühlen vorhanden sind, wird die universelle Anwendbarkeit des Auftritts als Metrik bei der Märchensegmentierung unumstritten.

Auftritt der handlungstragenden Figur im Märchen ist den im Drama explizit markierten Szenen gleichzustellen. Wenn wir Märchen als Drama-Stücke übertragen würden, könnte es im Prinzip mit dem in den Motiven segmentierten Text übereinstimmen (siehe unten die segmentierte Zusammenfassung des Typs ATU 551 - Das Wasser des Lebens).

Der Auftritt selbst kann im Text in verschiedenen Formen dargestellt werden, wie zum Beispiel:

1. Bericht über das, was ein Protagonist gerade erlebt oder worunter er leidet, zum Beispiel: Der König wird blind (König).
2. Bericht darüber, was gerade zwischen mehreren Protagonisten geschieht, zum Beispiel: Der König beauftragt seinen Sohn, nach seinem Tod sein Grab zu bewachen (König, jüngster Bruder, älteren Brüder).
3. Bericht darüber, was zwei Figuren gegeneinander reden, planen oder vorbereiten, ohne dass die dritte Figur aktiv beteiligt ist, zum Beispiel: Der jüngste Bruder gibt dem Freier seine Schwester zur Frau (Freier, jüngster Bruder, die Schwester kommt hier gar nicht zum Vorschein).
4. Bericht darüber, wie zwei Figuren gegenseitig interagieren, zum Beispiel: Der Jüngling tötet im Zweikampf das Ungeheuer (Jüngling, Ungeheuer).
5. Ein besonderer Fall ist die bei einem Auftritt erzählte Geschichte, die sogenannte Erzählung in der Erzählung, zum Beispiel: ATU 449 - Sidi Numan.

Die Liste lässt sich erweitern.

Den Motiven bevor sie in die Vollgeschichte integriert werden, lassen sich im mittleren Gebilde organisieren. Dieses Gebilde ist traditionell unter dem Namen Episode bekannt.

Durch den Auftritt einer handlungstragenden Figur im Märchen kann eine Episode gestartet, gefolgt oder beendet werden. Jede Episode stellt Einheit von folgenden Handlungsphasen dar:

– Krise auslösen) – (und) – Krise nachvollziehen – und – (Krise bekämpfen – oder – Krise missachten) – und – (Krise lösen – oder

Krise ist ein Zustand, der eine Lösung braucht. Jedes Motiv bezieht sich auf eine Krise und stellt eine Phase ihres Verlaufs dar.

Jedes Motiv ist nach seinem Wert und Sinn zu bestimmen. Nach dem Wert kann ein Motiv positiv, negativ oder ambivalent sein. Was den Sinn angeht, so kann ein Motiv entweder einen Bonus- oder Lastenausgang haben.

Ein Motiv mit Bonus-Ausgang dient der Förderung der handelnden Figur (es stellt ihr Ausrüstung, Gegenstände, Informationen, Versprechen oder Helfer zur Verfügung, die ihr in der Not helfen können). Ein Motiv mit Lastenausgang fordert hingegen den Verbrauch der gesammelten Bonus-Ressourcen, um die Behebung der Krise näher zu bringen.

Der Etalon für die Bestimmung vom Wert eines konkreten Motivs ist allein der Faktor des Helden und nicht der handelnden Figur. Die Teilnahme des Helden an einem Motiv macht es einfach, seinen Wert anhand der Folgen zu definieren. Problematisch sind Motive ohne Teilnahme des Helden. In solchen Fällen ist die Folge des Motivs von entscheidender Bedeutung. Hat das Motiv negative Folgen für den Helden, dann wird das Motiv als negativ bewertet. Hat das Motiv keine direkten Folgen für den Helden, dann wird der Wert des Motivs nach seinem absoluten Wert beurteilt. Scheitert ein falscher Held bei der Prüfung des potenziellen Stifters, handelt es sich um ein Bonus-Motiv mit negativem Wert (an der Stelle der handelnden Figur wäre dies für den Helden genau ein Bonus-Motiv mit negativen Folgen).

Um als offenkundige Krise betrachtet zu werden, muss diese entweder als ursprüngliche Krise existieren oder erst ausgelöst werden. Die Auslösung der Krise ist das Resultat einer negativen Handlung. Eine Krise existiert nicht als eigenständiges Motiv, sondern ist entweder die Folge einer Handlung oder die Offenkundigkeit eines solchen Handlungsergebnisses.

**Die Offenkundigkeit** der Krise ist die Situation, in der die Krise bekannt gemacht oder erkannt wird. Wie alles andere kann die Offenkundigkeit der Krise negativ, positiv oder ambivalent sein. Beispiele:

* Der Zwerg prüft die Höflichkeit der handelnden Figur, um ihr weiterzuhelfen. Hier liegt eine positive Absicht vor, die darauf abzielt, jemandem zu helfen.
* Ein Mädchen erkennt seinen Makel und wird neidisch auf die Schönheit der Schwester. Dies stellt eine negative Absicht dar, die darauf abzielt, jemanden zu verderben.

Der Grund für die Krise und die offenkundig gemachte Krise können manchmal ganz verkehrt erscheinen. Ein Beispiel: Der König fühlt sich durch den Helden bedroht (Grund der Krise) und will ihn beseitigen. Deshalb erteilt er dem Helden eine unlösbare Aufgabe, bei deren Erfüllung er sterben soll. Die Offenkundigkeit der Krise lautet jedoch nicht so, wie die Krise heißt, etwa: „Du musst vom Löwen gefressen werden“, sondern so: „Du musst ein Löwenfell beschaffen“. Weil die Offenkundigkeit das Ziel hat, den Helden zu verderben, ist sie trotz ihres Inhalts negativ. Dilemmatische Offenkundigkeit wäre z.B. ein Schild an der Kreuzung.

Der Kenntnisnahme der Krise wird eine entsprechende Handlung folgen. Eine positive Handlung schafft entweder einen Bonus für weitere Handlungen oder ermöglicht die Wiedergutmachung der Folgen einer negativen Handlung (Behebung der vorhandenen Krise). Eine negative Handlung lässt hingegen die vorhandene Krise bestehen oder löst eine neue Krise aus. Handlungen mit dilemmatischen Werten können beides tun, wie das Füttern des Vogels mit eigenem Fleisch oder die Tötung eigener Kinder zur Rettung des versteinerten Freundes.

Die Handlung kann verschoben werden. In solchen Fällen wird die Offenkundigkeit einer Krise durch die Offenkundigkeit einer nächsten Krise gefolgt. Es kann auch vorkommen, dass im Rahmen einer Krise geführte Handlungen durch eine neue Handlung aus dem Rahmen derselben Krise gefolgt werden. Gewöhnlich stellt eine solche Handlung einen Mehrwert der vorherigen Handlung dar, insbesondere in sogenannten Bonusepisoden. Obwohl der Mehrwert der Handlung als ein autonomes Motiv etabliert ist, ist er nach seinem Wert und Sinn von dem vorherigen Motiv nicht zu unterscheiden (daher auch sein Name – Mehrwert). Anders gesagt, das Resultat dessen, was und wie gehandelt wurde, ist vollkommen von der eigentlichen Handlung abhängig oder sogar in sie inkludiert. Ein Beispiel: Der ältere Bruder verhält sich frech gegenüber dem Zwerg und verpasst den Bonus von ihm. Das Verhalten des falschen Helden kann hier sowohl die Handlung als auch deren Ergebnis wahrgenommen werden.

Die Ergebnisse der Segmentierung des Märchens in Handlungsabschnitte lassen sich wie folgt zusammenfassen: Die kleinste empirisch erkennbare Entität für die strukturelle Segmentierung des Märchens entspricht dem Auftritt der handlungstragenden Person(en), d.h., der Szene. Der Wechsel zwischen den Szenen in der Erzählung entspricht dem Wechsel zwischen den Segmenten des Märchens. Die kleinste Kategorie dieser Segmente sind Motive, Motive bilden Episoden, die wiederum für die Bildung der Vollgeschichte zuständig sind. Ein Motiv ist entweder eine Offenkundigkeit oder der Handlung bzw. den Handlungen der in einer Episode thematisierten Krise gleich und beginnt und endet im Laufe eines Auftritts bzw. einer Szene.

Ein Auftritt besteht wenigstens aus einem Motiv und kann die Motive von mehr als einer Episode enthalten. Eine Episode setzt sich aus den Motiven zusammen, die Anfangs- und Lösungsphasen der in ihr thematisierten Krise widerspiegeln; das sind die Motive der Kategorien 'Offenkundigkeit' und 'Handlung der Krise'.

Zusammengefasst sind diese Kategorien wie folgt zu definieren:

1. Offenkundigkeit der Krise: Sich selbst oder anderen Handeln oder Nicht-Handeln ermöglichen (zum Handeln bringen).
2. Handlung der Krise: Sich allein oder mit fremder Hilfe, mit oder ohne Erfolg, mit einer Krise auseinandersetzen oder sich davon distanzieren.

Motive einer Episode können durch die Motive aus anderen Episoden getrennt werden, daraufhin kann eine Episode entstehen, die praktisch die ganze Geschichte umfasst. Die schematisierte Darstellung einer solchen Konstellation könnte wie folgt aussehen:

[Geschichte [Episode1 [Motiv1.1 + Motiv1.2]] + [Episode2 [Motiv2.1 + Motiv2.2 + Motiv2.3]] + [Episode1[Motiv1.3]]]

Wobei jedes Pluszeichen optional, und jede durch das Pluszeichen getrennte eckige Klammer bindend auf die Grenze zwischen zwei verschiedenen Szenen hinweist.

## Markup

### Tupel m

Das Markup wird uns helfen, Märchen so zu annotieren, dass alle ihre inhaltlichen Eigenschaften allein anhand dieser Annotationen vergleichbar sind. Die Entwicklung eines solchen Markups ist aufgrund der maximalen Verdichtung der inhaltlich-strukturellen Informationen möglich. Dabei dürfen die Anforderungen an die höhere Ausdruckspotenzial, Anwendungsfreundlichkeit sowie die Kompaktheit des Markups nicht aus den Augen verloren werden.

Zwecks der restlosen Erschließung der von uns oben erfassten Eigenschaften des Motivs (kurz ***m***) versuchen wir, es als ein Tupel von folgenden Attributen darzustellen:

Attribut ***a*** erfasst die Zugehörigkeit des Motivs zum Typ der Geschichte (ATU Index), ***b*** erfasst die Zugehörigkeit des Motivs zu einer Episodenphase sowie zum dessen Wert, ***c*** fasst den Inhalt des Motivs in einem Wortlaut zusammen, und ***d*** erfasst die Nomenklatur der im Motiv beteiligten Figuren.

Hier geben wir eine Beschreibung der Variationen jedes dieser Attribute im Einzelnen:

#### Attribut a

Jedes Motiv ist Teil einer Episode. Episoden bilden die Vollgeschichte. Also sind jede Episode und damit auch jedes Motiv Bestandteile eines Geschichtentyps. Diese sind, wie bekannt, in einem internationalen Typenkatalog erfasst und entsprechend indiziert. Wir bezeichnen diese Eigenschaft des Motivs mit dem Kleinbuchstaben ***a*** (für ATU) und der darauffolgenden Typennummer, zum Beispiel

*a300* für den Typ *ATU 300 – Drachentöter*.

Dieses Mal beschränken wir die Menge ***a*** auf die Indizes der Zaubermärchen.

#### Attribut b

Markierung der Zugehörigkeit des Motivs zu der bestimmten Episodenphasen ist durch folgende konventionellen Zeichen Möglich:

|  |  |
| --- | --- |
| Zeichenname | Die Episodenphase (Art des Motivs) |
| *Ef* | Offenkundigkeit der Krise |
| *Ha* | Handlung in der Krise |

Die Variationen nach dem unterschiedlichen Wert jeder von diesen Phasen lassen sich wie folgt unterscheiden:

|  |  |
| --- | --- |
| Zeichen | Der Wert der Episodenphase (Art des Motivs) |
| F (großes *Ef*) | Offenkundigkeit der Krise mit positivem Wert |
| f (kleines *Ef*) | Offenkundigkeit der Krise mit negativem Wert |
| Ff (beide *Ef*) | Offenkundigkeit der Krise mit ambivalentem Wert |
| H (großes *Ha*) | Handlung in der Krise mit positivem Wert |
| h (kleines *Ha*) | Handlung in der Krise mit negativem Wert |
| Hh (beide *Ha*) | Handlung in der Krise mit ambivalentem Wert |

Zusammengefasst sieht die Menge der Variationen des Elements *b* wie folgt aus:

#### Attribut c

Inhalt der Aktion stellt einen Wortlaut dar und zählt somit zum informativsten Bestandteil des Markupelements. Der Wortlaut entspricht dem inhaltlichen Wert des im Rahmen einer Episodenphase stattgefundener Aktion, z.B. "Heilmittel\_beschaffen".

Im Gegensatz zu anderen Attributen des Tupels, deren Variationen bereits vorbestimmt worden sind, ist eine endgültige Fertigstellung der Wortlautliste, also Variationen für das Attribut ***c****,* eher undenkbar. Es kann immer Variationen geben, die neue Formulierung fordern. Bereits aus diesem Grund zählt das Konzipieren eines Modells für die vorschriftliche Herstellung der Wortlautwerte zu wichtigste Aufgabe bei dem Bau des Markups. Ohne Lösung dieser Aufgabe wäre standardmäßiger Einsatz des Markups zum Scheitern verurteilt.

Durch die Nutzung der Tupel-Konstruktion des Markupelements ist es möglich ein und derselbe Wortlaut in bestimmten Fällen der Zusammenfassung des Inhalts mehr als einmal zu benutzen. Diese Fälle lassen sich wie folgt zu kategorisieren.

1. Klonen des Wortlauts
2. Ableitung eines Wortlauts aus dem anderen
3. Kombination von Klon- und Ableitungsmethoden.

Klonen oder unveränderte Anwendung eines Wortlautes ist gewöhnlich beim Wechsel zwischen der Offenkundigkeit als Beauftragung und dem Resultat der danach richtig oder falsch geleisteten Handlung möglich. Die Variation des Wortlauts ist dabei vorallererst durch den Wechsel von ***b*** Attribut zu erzielen. Die Manipulation durch die ***d*** Attributwerten ist außerdem möglich der Unterschied zwischen den durch den einen und denselben Wortlaut erfasste Inhalte weiter zu kontextualisieren.

Folgendes Beispiel:

*F:Heilmittel\_beschaffen:rHD\_rHH*

Das Element erfasst den vom Herrn des Helden zum Helden mit guter Absicht erteilten Auftrag, ihm ein dringend benötigtes Heilmittel zu beschaffen.

Durch den Wechsel ***F*** durch ***f*** und ***rHH*** durch ***rAN***, hätte derselben Wortlaut wesentlich anderen Inhalt zusammenfassen können. Das wäre nämlich der vom Antagonisten mit einer schlechten Absicht erteilte Auftrag, der Held bei der Erfüllung der unlösbaren Aufgabe hier, bei der Beschaffung des Heilmittels, verderben lassen (sog. Herkulesaufgabe). Das entsprechende Markupelement sieht wie folgt aus:

*f:Heilmittel\_beschaffen:HD\_rAN*

Durch den weiteren Wechsel der Werten vom Attribut ***b*** ist außerdem möglich der Mehrwert der positiv, negativ oder ambivalent ausgegangenen Handlung zu bezeichnen. Das folgende Markupelement kann z. B. wie folgt aufgelöst werden: Der Held endet seine Handlung zur Suche des Heilmittels mit positiven Folgen und händigt seinem Beauftragten das benötige Zielobjekt aus:

H:Heilmittel\_beschaffen:rHD\_rHH\_rZO

Kämme ein Zwischenfall vor, bei dem das richtige Lebenswasser von neidischen Brüdern (hier falschen Helden) durch das einfache ausgewechselt wird, so könnte auch diesen Fall durch denselben Wortlaut, mit gewechselten Werten von ***b*** und ***d*** Attributen erfasst werden:

*Hh:Heilmittel\_beschaffen:rHD\_rHH\_fZO*

Ein anderer Einsatz der Klonmethode ist für die Zusammenfassung des ersten Treffens zwischen dem Helden und seinem Stifter oder dem künftigen Weggefährten möglich. Das Letzte kann sowohl ein Helfer als auch ein falscher Held sein. Bei solchen Fällen handelt es sich gewöhnlich um die Probestellung der handelnden bzw. reisenden Figur und deren Reaktion. Die ältere Brüdern antworten frech auf die Frage des ihnen auf dem Weg begegneten Männleins uns und ernten Misserfolg bei der Erfüllung der Aufgabe. Der jüngste Bruder benimmt sich höflich gegenüber dem Männlein und sichert somit seine Hilfe, die in unterschiedlicher Form auftreten kann. Die Begegnung mit dem Zwerg und deren Vorgang kann man z. B. mit Hilfe folgendes Wortlauts erfassen *Höfliches\_Verhanten* Präfixe ***Ef*** und ***Ha*** hätten dabei auf die Phase des Treffens hinweisen können: *F:Höffliches\_Verhalten* als [verschleierte] Offenkundigkeit der Voraussetzung für die potenzielle Hilfe und *H:Höffliches\_Verhalten* als positive Reaktion des Helden auf diese Voraussetzung oder auch h:Höffliches\_Verhalten als negative, unhöfliche Reaktion des falschen Helden auf derselben Probe.

Im Gegensatz zu diesen Fällen lassen sich einige andere inhaltliche Konstrukte schwer oder gar nicht durch den gleichen Wortlaut erfassen. Am häufigsten ist das bei der Zusammenfassung der Auslösung und der Lösung der Krise der Fall, kann außerdem bei der Offenkundigkeit einer vorbestimmten Krise und deren Eintritt vorkommen. In solchen Fällen werden mehrere, jedoch aus einem Stichwort abgeleitete Wortlaute verwendet. Am häufigsten bilden solche Ableitungen ein Paar, z. B.

h:Schweigepflicht\_erhängen vs. H:Schweigenpflicht\_beheben

h:Kontakt\_abbrechen vs. H:Kontakt\_wiederherstellen

h:Auslösung\_von\_Krankheit vs. H:Krankheit\_Beheben

Oder sie können auch eine dreigliedrige Kette bilden:

h:Lebensgefahr\_Vergiftung > H:Handlungsanweisung\_Lebensgefahr\_beheben > H:Lebensgefahr\_beheben

Oft kommt es vor, dass der Zweite Mitglied des Paars einen weiteren Klon hat. Diese ist gewöhnlich die Offenkundigkeit der Krise, das später durch die Lösung der Krise gefolgt wird.

h: Auslösung\_von\_Krankheit\_ > F:Krankheit\_beheben > H:Krankheit\_beheben

Kommt solche Koppelung der Elemente im Spiel so reden wir um die Kombination von Klon- und Ableitungsmethoden Für der Erarbeitung des Markups.

Ein anderer, verhältnismäßig wesentlicher Teil des Markups besteht aus sog. Monoelemente. Sie bezeichnen gewöhnlich Ergebnisse oder Zwischenfälle der Handlungen, die durch die geklonten, abgeleiteten oder kombinierten Wortlaute erfasst sind. Z. B.:

H:Geboren\_werden

H:Eigene\_Lebensgeschichte\_Erzählen

H:Heiraten H:Bestrafung, H:Elixier\_erhalten usw.

Die Methoden wie Klonen und Ableitung der Wortlaute fördern einheitliche und freundliche Anwendung des daraus entwickelten Markups und garantiert die hohe Qualität der aus dessen Grund Annotierte Daten. Die einheitliche Anwendung der Markups wird außerdem durch den Wortlautkatalog unterstützt, der dem User ermöglicht die in der Wortlautliste Erfasste Werte gegenseitig vergleichbar und leicht auffindbar zu machen.

#### Attribut d

Wie bereits bekannt, kann es im Märchen insgesamt 24 handlungstragende Figuren geben. In einem Auftritt können gewöhnlich eins bis drei solchen Figuren agieren. Agieren bedeutet, dass die Figur in der Erzählung auftritt bzw. erscheint und nicht das, dass eine solche Figur von einer in der Episode agierende Figur erwähnt, gemeint oder bestellt wird).

Unter der Berücksichtigung der Höhenzahl der Kombination zwischen den Figuren, haben wir darauf verzichten müssen, die Reihenfolge der Figuren nach einer syntaktischen Funktion, wie z.B. das Subjekt / direktes Objekt / indirekte Objekt zu erfassen. Stattdessen haben wir überall die Interaktionen mit unterschiedlich kombinierten und/aber gleichen Figuren auf jeweils eine Kombination reduziert.

(ab|ba) = (ab)

(abc|acb|cab|aba|bac|bca) = (abc)

Dabei sind die handlungstragenden Figuren nach einer festen Reihenfolge (s. die Tabelle) kombiniert. D.h.: für die zweigliedrige Kombination wird wie folgt verfahren:

* Als erstens wird der rHD mit allen anderen kombiniert,
* Zweitens wird der rHH ebenfalls mit allen anderen kombiniert, ausgenommen der bereits eingesetzter rHD,
* Drittens wird das rRE mit allen anderen kombiniert, ausgenommen rHD und rHH usw.

Somit setzt sich die Menge der Variationen von ***d*** Elemente aus den Untermengen von Einzeln, zu zweit oder zu dritt agierenden Figuren zusammen:[[5]](#footnote-5)

{

*{rHD, rHH, rRE, rHF, rST rVB, rZM, rBZ, rZO, rHP, rPP, rAN, fAN, fPP, fHP, fZO, fBZ ... }*

*{rHD\_rHH, rHD\_rRE, rHD\_rHF, rHD\_ST, rHD\_rVB, rHD\_rZM, rHD\_rZO, rHD\_HP, rHD\_rPP ... }*

*{rHD\_rHH\_rRE, rHD\_rHH\_rHF, rHD\_rHH\_rST, rHD\_rHH\_rVB, rHD\_rHH\_rZM, rHD\_rHH\_rBZ ... }*

}

### Musterbeispiel

Im Folgenden versuchen wir, den Zusammenhang zwischen dem empirischen Objekt und dem darauf basierenden Modell anhand eines Beispiels zu veranschaulichen. Dazu werden wir einen klassischen Märchentext nach seinen Szenen segmentieren. [[6]](#footnote-6) Jedes Segment wird zusammengefasst und mit dem entsprechenden Markupelement versehen.

|  |  |  |  |
| --- | --- | --- | --- |
| 1. | Auftritt: |  | Der König |
|  | Inhalt: |  | Ein kranker König liegt im Sterben. |
|  | *Annotation:* | 1 | *a551h:Krankheit:rHH* |
| 2. | Auftritt: |  | Die drei Söhne und der alte Mann |
|  | Inhalt: |  | Ein weiser alter Mann weist die verzweifelten Königssöhne auf das schwer zu beschaffende Lebenswasser hin. |
|  | *Annotation:* | 2 | *a551:* *F:Heilmittel\_beschaffen:rHD\_rVB\_fHD* |
| 3. | Auftritt: |  | Der Vater und der älteste Sohn |
|  | Inhalt: |  | Der älteste Sohn bittet den Vater um Erlaubnis und macht sich auf den Weg, das Lebenswasser zu finden. |
|  | *Annotation:* | 3 | *a551:F: Werben\_um\_den\_Auftrag:rHH\_fHD* |
| 4. | Auftritt: |  | Der älteste Sohn und der Zwerg |
|  | Inhalt: |  | Der ältere Sohn zeigt Arroganz gegenüber dem Zwerg. Dieser wird zornig auf den ältesten Sohn. |
|  | *Annotation:* | 4  5 | *a551:F: Höfliches\_Verhalten:rST\_fHD*  *a551:h: Höfliches\_Verhalten:rST\_fHD* |
| 5. | Auftritt: |  | Der älteste Sohn und der Zwerg |
|  | Inhalt: |  | Der älteste Sohn wird in einer Felsspalte eingesperrt. |
|  | *Annotation:* | 6 | *a551:h: Festnahme\_bzw\_Sperren:rST\_fHD* |
| 6. | Auftritt: |  | Der Vater und der mittlere Sohn |
|  | Inhalt: |  | Der mittlere Sohn bittet den Vater um Erlaubnis und begibt sich auf die Suche nach Lebenswasser. |
|  | *Annotation:* | 7 | *a551:F:Werben\_um\_den\_Auftrag:rHH\_fHD* |
| 7. | Auftritt: |  | Der mittlere Sohn und der Zwerg |
|  | Inhalt: |  | Der mittlere Sohn zeigt ebenfalls Arroganz gegenüber dem Zwerg. Dieser wird wütend auf den mittleren Sohn. |
|  | *Annotation:* | 8  9 | *a551:F: Höfliches\_Verhalten:rST\_fHD*  *a551:h: Höfliches\_Verhalten:rST\_fHD* |
| 8. | Auftritt: |  | Der mittlere Sohn und der Zwerg |
|  | Inhalt: |  | Auch der mittlere Sohn wird in einer Felsspalte eingesperrt. |
|  | *Annotation:* | 10 | *a551:h: Festnahme\_bzw\_Sperren:rST\_fHD* |
| 9. | Auftritt: |  | Der Vater und der jüngste Sohn |
|  | Inhalt: |  | Der jüngste Sohn bittet den Vater um Erlaubnis und macht sich auf den Weg, Lebenswasser zu suchen. |
|  | *Annotation:* | 11 | *a551:F:Werben\_um\_den\_Auftrag:rHH\_fHD* |
| 10. | Auftritt: |  | Der jüngste Sohn und der Zwerg |
|  | Inhalt: |  | Der jüngste Sohn verhält sich höflich gegenüber dem Zwerg. Dieser gibt dem jüngsten Sohn notwendige Anweisungen. |
|  | *Annotation:* | 12  13  14  15 | *a551:F: Höfliches\_Verhalten:rHD\_rST*  *a551:H:Höfliches\_Verhalten: rHD\_rST*  *a551:H:* *Handlungsanweisungen\_Wache\_besänftigen:rHD\_rST*  *a551:h:Kontakt\_abbrechen:rHD\_rST* |
| 11. | Auftritt: |  | Der jüngste Sohn und die Wache |
|  | Inhalt: |  | Der jüngste Sohn besänftigt die Wache, die ihn passieren lässt. |
|  | *Annotation:* | 16  17 | *a551:H:Wache\_besänftigen:rHD\_fHF*  *a551:H:Wache\_entkommen:rHD\_fHF* |
| 12. | Auftritt: |  | Der jüngste Sohn und die verwunschenen Wesen |
|  | Inhalt: |  | Der jüngste Sohn nimmt die Zaubergegenstände. |
|  | *Annotation:* | 18 | *a551:H:Zaubermittel\_erhalten:rHD\_rZM* |
| 13. | Auftritt: |  | Der jüngste Sohn und die Besitzerin des Lebenswassers |
|  | Inhalt: |  | Es entwickelt sich eine Liebesbeziehung zwischen dem jüngsten Sohn und der Besitzerin des Lebenswassers. |
|  | *Annotation:* | 19  20 | *a551:F:* *Handlungsanweisung\_Objekt\_ergreifen:rHD\_fBZ*  *a551:h:Kontakr\_abbrechen: rHD\_fBZ* |
| 14. | Auftritt: |  | Der jüngste Sohn und das Lebenswasser |
|  | Inhalt: |  | Der jüngste Sohn erlangt das Lebenswasser. |
|  | *Annotation:* | 21 | *a551:H:* *Objekt\_ergreifen:rHD\_rZO* |
| 15. | Auftritt: |  | Der jüngste Sohn und der Zwerg |
|  | Inhalt: |  | Der jüngste Sohn bittet den Zwerg um die Freilassung seiner Brüder, ignoriert dabei aber dessen Warnung. |
|  | *Annotation:* | 22  23  24 | *a551:F:* *Kontakt\_wiederherstellen:rHD\_rST*  *a551:F:* *Verrat\_vermeiden:rHD\_rST*  *a551:h:* *Verrat\_vermeiden\_missachten:rHD\_rST* |
| 16. | Auftritt: |  | Der jüngste Sohn und seine Brüder |
|  | Inhalt: |  | Der jüngste Sohn trifft seine befreiten Brüder und teilt ihnen von seinem Erfolg mit. |
|  | *Annotation:* | 25 | *a551:Hh:Festnahme\_bzw\_Sperre\_beheben:rHD\_fHD* |
| 17. | Auftritt: |  | Die Könige aus fernen Ländern und der jüngste Sohn |
|  | Inhalt: |  | Auf dem Rückweg nach Hause unterstützt der jüngste Bruder die Herrscher von drei unterschiedlichen Ländern mithilfe seiner magischen Gegenstände. |
|  | *Annotation:* | 26  27  28 | *a551:F:Rettung\_der\_Hilfsbedürftigen:rHD\_rHF\_rZM*  *a551:H:Rettung\_der\_Hilfsbedürftigen:rHD\_rHF\_rZM*  *a551:h:Kontakt\_abbrechen:rHD\_rHF* |
| 18. | Auftritt: |  | Die Könige aus fernen Ländern und der jüngste Sohn |
|  | Inhalt: |  | Auf dem Rückweg nach Hause unterstützt der jüngste Bruder die Herrscher von drei unterschiedlichen Ländern mithilfe seiner magischen Gegenstände. |
|  | *Annotation:* | 29  30  31 | *a551:F:Rettung\_der\_Hilfsbedürftigen:rHD\_rHF\_rZM*  *a551:H:Rettung\_der\_Hilfsbedürftigen:rHD\_rHF\_rZM*  *a551:h:Kontakt\_abbrechen:rHD\_rHF* |
| 19. | Auftritt: |  | Die älteren Söhne beim Schlaf ihres jüngsten Bruders |
|  | Inhalt: |  | Die älteren Brüder werden neidisch und tauschen heimlich das Lebenswasser gegen gewöhnliches Wasser aus. |
|  | *Annotation:* | 32  33 | *A551:f:* *Missetat\_planen:rZM\_fHD*  *a551:h:* *Missetat\_Aneignen:rZM\_fHD* |
| 20. | Auftritt: |  | Der jüngste Sohn und ihr Vater |
|  | Inhalt: |  | Das beschaffte Heilmittel macht den König noch kränker. |
|  | *Annotation:* | 34  35 | *a551:* *h:Heilmittel\_beschaffen:* *:rHD\_rHH\_fZO*  *a551:h:Krankheit\_beheben:rHD\_rHH\_fZO* |
| 21. | Auftritt: |  | Die älteren Söhne der Vater und das Heilmittel |
|  | Inhalt: |  | Die älteren Söhne heilen den Vater und verleumden den jüngsten Bruder. |
|  | *Annotation:* | 36  37  38 | *a551:* *h:Heilmittel\_beschaffen:* *:* *rHH\_rZO\_fHD*  *a551:Hh:Krankheit\_beheben:* *rHH\_rZO\_fHD*  *a551:h:Verleumdung\_von\_Untreue:* *rHH\_rZO\_fHD* |
| 22. | Auftritt: |  | Die älteren und der jüngste Bruder |
|  | Inhalt: |  | Die älteren Brüder bringen den jüngsten Bruder zum Schweigen. |
|  | *Annotation:* | 39 | *a551:h: Zum\_Schweigen\_bringen:rHD\_fHD* |
| 23. | Auftritt: |  | Der König und der Scharfrichter |
|  | Inhalt: |  | Der König befiehlt den Tod seines jüngsten Sohns. |
|  | *Annotation:* | 40 | *a551:f:Todesauftrag:rVB\_fHH* |
| 24. | Auftritt: |  | Der ältere Sohn und der Scharfrichter |
|  | Inhalt: |  | Der Scharfrichter weigert sich, den Prinzen zu töten, und dieser flieht. |
|  | *Annotation:* | 41  42 | *A551:H:* *Auslösen\_von\_Mitleidsgefühl:rHD\_rVB*  *a551:H:* *Todesauftrag\_entkommen:rHD\_rVB* |
| 25. | Auftritt: |  | Der Vater, die drei fremden Könige |
|  | Inhalt: |  | Die drei Könige, denen der Prinz einst half, besuchen das Königreich. Der König erfährt die Wahrheit und vergibt seinem jüngsten Sohn. |
|  | *Annotation:* | 43  44 | *a551:H:Kontakt\_wiederherstellen:rHH\_rHF*  *a551:H:* *Missetat\_enthüllen: rHH\_ rHF* |
| 26. | Auftritt: |  | Der König und der Scharfrichter |
|  | Inhalt: |  | Der König erfährt vom Scharfrichter, dass sein jüngster Sohn freigelassen wurde, und erkennt die Verleumdung. |
|  | *Annotation:* | 45 | *a551: H:Verleumdung\_beheben:rHH\_rVB* |
| 27. | Auftritt: |  | Die Besitzerin des Zaubermittels |
|  | Inhalt: |  | Gleichzeitig bereitet sich die Besitzerin des Lebenswassers auf die Suche nach ihrem Geliebten vor. |
|  | *Annotation:* | 46 | *a551:F:Kontakt\_wiederherstellen:fBZ* |
| 28. | Auftritt: |  | Der älteste Sohn und der Diener der Besitzerin des Lebenswassers |
|  | Inhalt: |  | Der älteste Sohn scheitert ebenfalls beim Identitätstest. |
|  | *Annotation:* | 47 | *a551:h:Identität\_testen:fHD* |
| 29. | Auftritt: |  | Der mittlere Sohn und der Diener der Besitzerin des Lebenswassers |
|  | Inhalt: |  | Der mittlere Sohn scheitert ebenfalls beim Identitätstest. |
|  | *Annotation:* | 48 | *a551:h:Identität\_testen:fHD* |
| 30. | Auftritt: |  | Der jüngste Sohn und die Besitzerin des Lebenswassers |
|  | Inhalt: |  | Der jüngste Sohn besteht den Test und wird erkannt. |
|  | *Annotation:* | 49  50 | *a551:H:Identität\_testen:rHD\_fBZ*  *a551:H:Kontakt\_wiederherstellen:rHD\_fBZ* |
| 31. | Auftritt: |  | Vater und der älteste Sohn |
|  | Inhalt: |  | Der älteste Sohn gesteht dem König die Wahrheit. Die Verbrecher fliehen und kehren nie zurück. |
|  | *Annotation:* | 51  52 | *a551:H: Schweigenpflicht\_beheben:rHD\_rHH\_fHD*  *a551:H:Bestrafung:rHD\_rHH\_fHD* |

*.*

## Motiv, Episode, Märchen

Wir können die Organisation der inhaltlichen Elemente im Märchen auf zwei Strukturebene bestätigen: das sind die Motive und die Vollgeschichte. Das erste lässt sich durch den Anfang und Ende der Szenen in der Geschichte erkennen, das andere durch den Anfang und Ende der gesamten Geschichte bzw. des Textes.

Die Vollgeschichte als oberste Instanz der Analyse entspricht einer Makrostruktur, die aus den in den mesoskopischen Strukturelementen organisierten Mikroelementen besteht. Bei der Annahme dieser Definition für die Vollgeschichte ist das Motiv hingegen als kleinstes und damit als mikrostrukturelles Element des Systems zu definieren. Was zwischen der Vollgeschichte und dem Motiv vorkommt, gehört zum mesoskopischen Strukturelement. Als solches halten wir die Episode möglich.

Die mesoskopische Natur der Episode beinhaltet automatisch die Repräsentationsmöglichkeit der Charakter von Mikro- und Makroelemente durch das mesoskopischen Element. Daher ergibt sich die Möglichkeit, dass eine Episode sowohl ein Baustein als auch ein Gebilde sein kann, das selbst aus solchen Bausteinen besteht.

Anhand unseres Musterbeispiels versuchen wir den Zusammenhang zwischen den Strukturelemente ***Motiv > Episode > Märchen*** zu präsentieren. Dafür werden wir die Annotation des Textes als Motivliste darstellen (sie die Werte *Annotation* in der Tabelle 2) und die Zusammenhänge der Motive in den Episoden erfassen.

In einer Episode wird die Geschichte einer großen oder kleinen Krise erzählt. Diese Krisen sind durch die direkte oder indirekte Teilnahme des Helden vollzogen. Für die Definition einer Krise ist deshalb wichtig diese nicht nur anhand der Handlung sondern auch in ihr agierenden Figuren zu charakterisieren.

oder beziehen sich indirekt Die Episode folgt einer Struktur, die die Auslösung, Handlung und Lösung dieser Krise behandelt. Wie bereits bekannt, wird die Lösung der Krise durch das Motiv mit den Attributwerten ***b=H/Hh*** dargestellt. Die Lösung der Krise wird hingegen durch ein Motiv erfasst, das dem letzten vorabgeht, keinen Attributwert ***b=H*** hat und einen vergleichbaren Wortlaut (Attributwert ***c***[[7]](#footnote-7)) besitzt.

Andere Motive, die sich auf dieselbe Krise beziehen, sollen sich in unmittelbarer Nähe zu den oben genannten Motiven befinden und gleiche Attributwert ***d*** haben. Dazu können auch Motive gehören, die durch ihre Attributwert ***c*** mit dem letzten vergleichbar sind.

Der gegenseitige Vergleich der in der Liste erfasste Motive läuft von unten nach oben. Wird ein Motiv als Bestandteil einer Episode erfasst, darf es nicht mehr für die Bildung einer anderen Episode verwendet werden.

Somit lassen sich 8 Episoden in der gesamten Geschichte ausmachen. Jede einzelne Episode von denen hat ihr eigenes Thema und eigenen Kreis von handlungstragenden Figuren.

* 1. Krankheit und die Heilung: 1-2, 20, 30-34
  2. Verhalten gegenüber dem Zwerg mit zweierlei Folgen: 4-6, 8-10, 12-16, 21-23
  3. Beziehung mit der Besitzerin des Zielobjektes: 18,19, 45-49
  4. Rettung der künftigen Zeugen: 24-27
  5. Missetat von älteren Brüdern: 28, 29, 39,40
  6. Verleumdung
  7. und die Rolle von drei Königen: 21-23, 25 -32, 42-43
  8. Todesauftrag: 39-41

Motiven wie 3,7,11 und 18 werden nicht erfasst. Sie

# Maschinelle Analyse des Märchens

Formalisierte Beschreibung der Struktur des Märchens eröffnet den Weg zu maschineller Bearbeitung des Forschungsobjektes. Diese meint Automatisierung sowohl die Segmentation des Märchens in Strukturelemente von unterschiedlicher Größe als auch Erzeugen eines zusammenfassenden Bildes aufgrund des Vergleichs der je nach der Fragestellung ausgewählte Forschungsobjekten, wie etwa: vergleichende Analyse des Typs X anhand der Repertoires Y und Z. Kombination von Motiven innerhalb der Episode X. Status der Handlungstragende Figuren im Motiv Y und vieles andere.

Die Lösung der Maschinelle Analyse des Märchens als nächster Schritt nach der Formalisierung des Forschungsobjektes ist nur mit der Entwicklung der entsprechenden digitale Infrastruktur verbunden. Im Folgenden werden wir einige wichtigste Elemente dieser Infrastruktur betrachten.

## Textkorpus

Das Textkorpus stellt eine auf dem TEI-Standard basierte XML-Datei dar. Als Wurzelelement der Datei gilt <teiCorpus>. Der Korpus kann Texten in verschiedenen Sprachen auffangen. Als einzige Voraussetzung dabei gilt die standardisierte Kodierung der Texte in Unicode und ihre Konformität mit der Open-Data Forderungen. Sollen sich die Texte nicht an die letzte Voraussetzung halten, kann man sie trotzdem in der abgeleiteten Form verwenden. Als solche könnte die von Stoppwörtern bereinigte Texte gelten. Der gezielten Zugriff auf die Texte aus verschieden Sprach oder/und anderen Gruppe kann man anhand der im <teiHeader> - Element erfassten Metadaten bewerkstelligen, als einfacher gilt dabei die Anwendung des xml:id Attributs beim Element <Text>. Dieses Attribut, das aus drei Teilen besteht, kodiert die Information über folgenden Eigenschaften des Textes:

* 1. die Zugehörigkeit des Textes zu bestimmtem Repertoire
  2. die Sprache des Textes
  3. laufende Nummer des Textes unter den anderen Materialien mit gleicher Zugehörigkeit zum Repertoire.

Folgende Beispiele

* *deu\_deu\_1* = 1. Text aus dem deutschen Repertoire in deutsche Sprache
* *deu\_eng\_1* = Derselbe 1. Text aus dem deutschen Repertoire in englische Sprache
* *deu\_eng\_2* = 2. Text aus dem deutschen Repertoire in englische Sprache usw.

Für die Bezeichnung der Namen der Repertoires und Sprachen wird die Abkürzungen laut dem Standard ISO 639-3 verwendet.

Im TEI-Korpus sind die Märchen einzeln in einem separaten TEI-Element erfasst. Für jedes Märchen gibt es eine codierte, bzw. für die Codierung bereitgestellte Textfassung. Die Basistexte befinden sich in verschiedene Repositorien und sind miteinander und dem codierten Text entsprechend referenziert.

Die nicht rechtkonforme Texte sind im Textkorpus in abgeleitete Form dargestellt und mit bibliographischen Angaben über die gedruckte Textquelle versehen.

Die Hauptaufgabe des Korpus besteht darin eine standardisierte Segmentation der Daten sowie einen gezielten Zugriff in diesen Daten zu gewährleisten. Das zu diesem Zweck entwickelte Datenmodell macht möglich die im Text vorhandene Szene auszumachen und diese ihren Bestandteilen nach zu erfassen. Dafür ist im Element *<seg>* eingebettete Tupel von ***a,b,c*** und ***d*** Attribute zuständig. Als optimale Anzahl von Motive innerhalb einer Szene beträgt fünf. Dies bedeutet nämlich die Erfassungsmöglichkeit jeweils drei ***a, b,*** und ***c*** Attribute und eine für alle fünf Tupel gemeinsames ***d*** Attributes. In unserem Text wäre z.B. der 10. Abschnitt wie folgt annotiert:

*<seg a1:ana = “a551“ b1:ana = “F“ c1:ana = “ Höfliches\_Verhalten“ b2:ana = “H“ c2:ana = “ Höfliches\_Verhalten“ b3:ana=“H“ c3:ana=“ Handlungsanweisungen\_Wache\_besänftigen“ b4:ana=“0“ c4:ana=“ 0“ b5:ana=“0“ c5:ana=“ 0“ d:ana=“ rHD\_rHF“ >* Die Begegnungsszene zwischen dem Held und dem Stifter *</seg>*

## Künstlicher Assistent für die semiautomatische Etikettierung der Texte

Zu den wichtigsten Komponenten der maschinelle Analyse der Märchen gehört der Assistent mit Hervorsagefunktion. Er soll dafür zuständig sein in unüberschaubare Menge von Daten die Textteile zu erkennen die über die von uns gesuchte inhaltliche Eigenschaften verfügen.

Entwicklung eines solchen Assistenten ist auf die Erfüllung folgender drei Aufgaben angewiesen:

1. unter der vorhandenen Vorgehensweise der maschinellen Analyse des Textes ein passendes auszuwählen
2. (ii) diese mit dem Forschungsfrage anzupassen
3. (iii) das gewonnene Wissen über das Forschungsobjekt in Sprache der ausgewählten Vorgehensweise umzuformulieren.

Im Rahmen der letzten Aufgabe ist insbesondere die Anpassung zwischen den konzeptuellen und empirisch beobachtbaren Elementen hervorzuheben. Wie oben angedeutet gilt eine Szene innerhalb der Geschichte als empirisch beobachtbare Grenze für den Anfang bzw. das Ende des Motives. Für die Maschine bedeutet das eine Zeichenkette deren kleinste erkennbares Element ein Wort und das größte ein Absatz eher unwahrscheinlich mehrere Absätze sind. Auf der inhaltlichen Ebene entspricht dieses dem Textteil mit geänderten Figuren, ihren Attributen, Ortschaften und Handlungen. Wir können Maschine beibringen diesen Wechsel zu erkennen, aber dafür muss erfasst werden welcher Zusammenhang zwischen den Strukturen von Roh-, Training-, Mess- und Zieldaten besteht.

Der rohe Text so, wie er uns gewöhnlich vorliegt ist durch die folgenden explizit erkennbaren Zeichen wahrnehmbar: Wort, Satz, Absatz und Text. Die Grenzen der somit organisierten Textteile entsprechen nicht unabwendbar dem von uns zu suchende Szenenwechsel im Text.

Die Aufbereitung des rohen Textes zu Trainingsdaten ist deshalb auf die Umwandlung der sog. *Istzeichen* in *Sollzeichen* angewiesen. Dafür gilt am häufigsten die Zahl der Absätze im Text zu korrigieren. Sie können sowohl zusammengelegt als auch in mehreren Teilen geteilt werden. In manchen Fällen können einige solche Absätze auch einem einzigen Satz gleich sein.

Relativ einfacher ist die Erkennung der Texte mit spezifischen Episoden bzw. Episodengruppe. Dabei bezieht sich der Assistent auf den gesamten Text und stellt fest in welchen von der gesamten Menge der Texte die Gesuchte Episode zu finden sind.

Für die automatischen Erkennung der Motive Absätzen und Episoden im Text legen wir den Algorithmus der logistischer Regression *(LogisticRegression)* zugrunde.[[8]](#footnote-8) Ein durch diesen Algorithmus gesteuertes Modell versucht die für die Erkennung bzw. das Extrahieren der gesuchten Episode (positiv etikettierte Datensätze) und ihre Gegenteile (negativ etikettierte Datensätze) anwendbaren Merkmale zu finden und sie auf der einen und andere Seiten der imaginären Trennlinien anzuordnen.

Entscheidend für die Einnahme der Elemente im Modell gelten die in den TF-IDF Werte skalierten Merkmalen aus positiv und negativ ausgewerteten Datensetzen (diese werden im Vorfeld aus der sog. Stopp- und andre für den Vergleich irrelevante Wörter bereinigt).

Aufgrund der gewöhnlich anzutreffenden mäßigen Anzahl von positiv etikettierten Trainingsdaten wird eine mehrfache Kreuzvalidierungsmethode eingesetzt. Sie sorgt dafür, dass die gesamten Trainingsdaten in gleichmäßige Abschnitte geteilt und als Validierungs-, bzw. Trainingsdaten verwendet werden. Das so trainierte Modell sollte nun im Stande sein, die für die positiv und negativ etikettierten Daten charakteristischen Merkmalen (Wortformen) trennen und sie als eine rangierte Liste der entsprechenden Zeichengruppen darstellen zu können.

Das trainierte Modell wird durch die Vorhersagefunktion erweitert.[[9]](#footnote-9) In unserem Fall scheitert zwar die Funktion bei einigen Vorhersagen, sie kann dennoch als zuverlässiges Werkzeug gelten. Der genaue Grund, warum die Hervorsagefunktion bei einigen Fällen scheitert, ist nicht eindeutig zu ermitteln. Es ist allerdings anzunehmen, dass einiges an der Abgleichsmethode der Ähnlichkeit zwischen dem Modell und den Messdaten liegen könnte. Die Sigmoid Funktion, die bei der Berechnung der Distanz zwischen dem Modell und den Messdaten mit Erfolg eingesetzt werden kann, könnte hier dadurch verhindert werden, dass der Datensatz oft mit beiden Klassen (sowohl positiv als auch negativ) dargestellt wird. Die angenommene Ähnlichkeit mit dem gesuchten Inhalt bzw. der Episode verläuft in solchen Fällen womöglich fehlerhaft, weil negative und positive Merkmale hier gegenseitig wirken.

Unter der Berücksichtigung des obengenannten Faktors scheint es viel effektiver zu sein, dass man die Daten ausschließlich nur anhand der mit höchsten Koeffizienten versehenen positiven Merkmale untersucht. Damit werden die Datensätzen nicht über das Modell verglichen, sondern über das aus dem Modell extrahierte prototypische Metamotiv. Diese funktioniert wie eine Suchabfrage an die Dokumentensammlung.

Ein Bild, das Screenshot, Reihe, Rechteck, Diagramm enthält.

Automatisch generierte Beschreibung

Abb. 6

Das Extrahieren der Merkmale erfolgt durch die Berechnung des Schwellenwertes für die Topklassifikationsmerkmale. Dieser Wert ist an der Stelle der Merkmal-Kette zu suchen, wo die Gewichte der positiven und negativen Merkmale bzw. Koeffizienten abgeglichen werden.

Nach dem graphischen Verfahren kann man diese Stelle wie folgt lokalisieren: Die Abbildung 6 zeigt die Schwingung zwischen den maximalen und minimalen Koeffizienten der negativ und positiv eingeschätzten Klassifikationsmerkmale. Die für beiden Kategorien geltende maximale Zahl der Merkmale (z) ist eine Variable und kann entsprechend der Situation angepasst werden, diesmal beträgt sie 200. Blaue Spalten stellen den positiven Merkmalen dar, die roten Spalten visualisieren hingegen die negativen Merkmale (Wegen der hohen Anzahl überlappen sich die Merkmale (siehe schwarze Wortwolke im unteren Bereich der Grafik). Für die Visualisierung der gesuchten Stelle wird die vorhandene Grafik an dem Treffpunkt von roten und blauen Spalten geteilt und die linke Hälfte so weit nach rechts verschoben, bis die kleinste rote Spalte unter der größten blauen Spalte platziert wird. Beachtet man die Größe von blauen und roten Spalten kann man leicht feststellen, an welcher Stelle sie übereinstimmen (s. Abb. 6.1).

Ohne graphische Darstellung kann man derselben Schwellenwert anhand folgender Formel berechnen:

Kommt die Gleichung zustande:

, *wobei ,*

so ist der Schwellenwert gleich *.*

Da die Topklassifikationsmerkmale nichts anderes darstellen als eine einfache Liste von positiv bewerteten Elementen bzw. Wörtern, ist ihre Bildung mit Hilfe einer einfachen Stufenfunktion möglich (entspricht der grünen Grafik auf der Abb. 6.1):

Ein Bild, das Screenshot, Reihe, Rechteck, parallel enthält.

Automatisch generierte Beschreibung

Abb. 6.1

Die Liste der auf diese Weise festgelegten Merkmale wird als prototypisches Metamotiv in einem Datensatz gespeichert und bei den Messdaten angelegt.

Die Erkennung des gesuchten Motivs in Messdaten erfolgt durch die Klassifikation auf der Basis der Kosinus-Ähnlichkeit.[[10]](#footnote-10) Diese meint den Vergleich jedes Datensatzes mit dem synthetischen Datensatz (dem sog. prototypischen Metamotiv), das seinerseits aus den gewonnenen Klassifikationsmerkmalen des Logistic Regression Modells besteht.

Die Kosinus-Ähnlichkeit zwischen *a* und *b* Datensätzen wird nach folgender Formel berechnet:

Dabei entspricht ein *ai*, bzw. *bi* einem, in einem TF-IDF Wert skalierten Merkmal (Token) aus einem *a*, bzw. *b* Datensatz (Vektor). Je näher der Wert der berechneten Cosinus-Ähnlichkeit bei Eins liegt, desto plausibler ist die vorausgesagte inhaltliche Ähnlichkeit zwischen zwei (*a* und *b*) Datensätzen.

## Tools für die Annotation und Auswertung der Daten

Durch die inhaltliche Annotation wird es möglich, eine große Menge von Texten einheitlich und mit präziser Fokussierung auf die konkreten Forschungsfragen auszuwerten und Ergebnisse nachvollziehbar und anschaulich darzustellen. Hier betrachten wir Anwendungsbeispiel mit dem Ziel die Normalform eines Typs in Westeuropäische Kulturraum in der visualisierten Form auszuwerten.

Im Vorfeld der Auswertung muss es vorbestimmt werden (a) nach welchem Typ bzw. Typencluster wir suchen und (b) im Welchem Teil der Daten die Suche durchgeführt werden soll.

Angenommen, interessieren wir uns für die Normalform des Märchentypus *a551 – Wasser des Lebens* im europäischen Märchenrepertoire[[11]](#footnote-11) (Deutsch, Österreichisch, Französisch, Italienisch, Ungarisch, Griechisch, Russisch).

Nach der Eingabe von entsprechenden Suchparameter greift der Assistent auf die annotierten Märchentexten mit passenden Identifikatoren zu und stellt die Auswertungsergebnisse in der Form von Graphen und dabei in der einfachen Liste der Kompositionsbau beteiligten Motive dar.

Der Graph zeigt ein zusammengeführtes Bild der in den verschiedenen Texten gegenseitig geketteten Motiven mit gesuchtem ***a*** Attributwert. An den Peripherien der Kette können auch einige Motiven auftreten, die andere ***a*** Attributwerte haben. Diese stellen ihrerseits die Peripherien der Motivkette dar, die den mit dem gesuchten Typ kombinierten Typen ausmachen.[[12]](#footnote-12) Nach der Eingabe der Werten von solchen ***a*** Attribute ist zu erwarten, dass der Graph der Beziehungen von Motiven innerhalb eines Typs zum Graphen der Beziehungen der Motive innerhalb eines Typenclusters umgewandelt wird.

Die Auswertung der Motivbestand in einem Typ bzw. Typcluster, wird automatisch durch die Aggregation der Bibliothek der dabei involvierten Motive gefolgt. Diese kann, solange es keine entsprechende Aufforderung vorliegt, eine einfache Reihe von Motiven ohne Hinweise auf jegliche Verbindung darstellen. Bei der Eingabe eines konkreten Motivs wird die Liste wie folgt reorganisiert: das eingegebene Motiv bekommt die Nummer Null und wird an seine Stelle in der Reihe von gesamten Motiven platziert, die auf der Zeitachse ihm nach vorne und nach hinter liegende Motive bekommen entsprechende Nummerierung mit Minus- und Pluszeichen und werden den Platz oberhalb bzw. unterhalb des mit Null bezeichneten Motivs übernehmen. Wird die Kette der Motive von einem Typ durch eine oder mehrere Motiven mit einem anderen ***a*** Attribut getrennt so werden sie ohne Hinweis auf ihre Position in der Motivkette erfasst. Somit erhalten wir eine Liste von Motiven, die im dem zusammengefassten Inhalt des ausgewerteten Typs übereinstimmt. Der wissenschaftliche Wert und Plausibilität einer solche Auswertung hängt vollkommen mit der repräsentativitätsgrad der erfassten Forschungsdaten.

erfasst die in unmittelbaren Kontakt mit so wie sie in ihren Herkunftstexten in bestimmten Reihenfolgen miteinander verkettet sind. Das Skript kann alle diese Reihenfolgen zusammenführen und in einer Multikette von Episoden des gesuchten Typs und der mit ihnen verknüpften eine oder mehrere Episoden (aus dem gesuchten oder einem fremden Typ) darstellen.

Dafür müssen wir mit einem Extra-Eintrag eine Episodengruppe festlegen und an einer bestimmten Stelle im Skript notieren.[[13]](#footnote-13) Für unseren Fall wählen wir die Episodengruppe *e301\_*.

~~Die erzeugte Liste besteht aus Tripletten (s. Abb. 7, mit doppelten Linien getrennten Zeilengruppen). Jedes Triplett besteht aus einer mittleren Episode (s. Abb. 7, mit einfachen Linien umgebene Zeilen), welche mit ein bzw. mehreren vorhergehenden bzw. folgenden Episoden umgeben wird. Jedes Glied aus der gesuchten Episodengruppe soll als Triplett zugleich alle drei Glieder der Episodengruppe darstellen. Hat eine Episode kein Anfangs- oder Ergänzungsglied, so werden an entsprechender Stelle synthetische Anfangs- bzw. Endglieder jeweils einer Episode in der Liste generiert: diese bestehen aus der Abkürzung ‚Anf-‘ bzw. ‚End-‘ und dem Präfix der Episode. Maschinell wird dadurch jede [echte] Episode des gesuchten Typs als mittleres Bindeglied der Episodenkette wahrgenommen und somit als vorläufige oder/und nachfolgende in Bezug auf jeweils eine andere Episode betrachtet. Für den Mensch ist das ein Hinweis auf den Status der Episode, die in der Episodenkette eines Texts entweder an der ersten oder an der letzten Position steht.~~

~~Jede Episode in der Liste wird mit einem Häufigkeitsindex versehen. Die Liste zeigt somit, (i) durch welche Episoden der Typ am häufigsten realisiert wird (Reihenfolge der Mittel-Episoden in der Liste) und (ii) mit welchen Episoden aus eigenen oder anderen Kategorien sie am häufigsten kombiniert sind (Bezug der mittleren Episode auf die vorhergehenden und folgenden Episoden in der Liste).~~

Weitere synoptische Darstellungen der inhaltlichen Struktur des Typs in einem Repertoire erfolgen durch die Visualisierung der oben beschriebenen Liste durch die Graphen.

Im Gegensatz zu der Liste stellen die Graphen jede Episode als ein Knoten und jede Verbindung zweier Knoten als Kante dar.

Abfolge der Handlung: Als Anfangsepisode für den Typ a301 gelten e301\_c\_, e650A\_e, e650A\_c und e650A\_m\_. Jede einzelne von diesen Episoden kann den Beginn der Episode e301\_e vorbereiten. Eine andere Möglichkeit wäre der Beginn des Typs direkt von der Episode e301\_e\_. Nach dem Graph endet die gesamte Geschichte in den meisten Fällen mit der Episode e301\_o\_, selten kann sie aber bereits bei der Episode e301\_i\_ abgebrochen werden.

Der Graph erfasst die Realisierung des Typs a301 vollständig (insofern dies das Material erlaubt). Den am häufigsten anzutreffenden Episoden nach (hier die mit den roten Kanten verbundenen Knoten), ist die Normalform des Typs 301 wie folgt zusammenzufassen:

* e301\_e\_vom\_beginn\_der\_aktivitaet\_bis\_zum\_finden\_der\_raststaette\_oder\_eines\_eingangs

Der Held geht in weite Welt. Unterwegs gewinnt er mehrere Weggefährten mit außergewöhnlichen Eigenschaften. Nach einiger Zeit kommen die Wandernden an einen Ort, wo sie ihre provisorischen Lager aufschlagen.

* e301\_g\_vom\_finder\_einer\_nachtrast\_bis\_zum\_betreten\_der\_welt\_des\_AN

Am nächsten Tag übernimmt einer den Dienst in der Küche, wenn die anderen auf die Jagd gehen. Er wird von einem Männchen angegriffen und drangsaliert. Das wiederholt sich auch bei dem anderen Weggefährten. Im Gegensatz zu seinen Kameraden überwältigt der Held das Männlein. Es flieht und verrät somit den Eingang zu der Unterwelt. Der Held lässt sich abseilen und gelangt zu der Behausung eines Ungeheuers.

* e301\_i\_vom\_betreten\_des\_wohnsitzes\_des\_AN\_bis\_zur\_befreiung\_des\_OB

Der Held tötet das Ungeheuer im Zweikampf und befreit die von diesem entführte Prinzessin (manchmal drei).

* *e301\_k\_empfang\_des\_OB\_und\_verrat\_des\_FH*

Die oben gebliebenen Kameraden ziehen die Prinzessin hoch und lassen den Helden im Abgrund zurück.

* e301\_m\_entkommen\_aus\_dem\_haus\_des\_AN

Der Held gelangt an die Oberwelt mit Hilfe desselben Männleins, eines Geistes, der ihm Flugkraft verschafft oder eines Vogels, den er beim Flug mit seinem eigenen Fleisch füttert.

* e301\_o\_von\_der\_rueckkehr\_an\_die\_oberflaeche\_bis\_zur\_loesung\_der\_durch\_den\_raub\_verursachten\_krise

Genau am Tag der Hochzeit kehrt der Held zurück und löst mit Gewalt die Zwangsehe zwischen Prinzessin und einem der Betrüger auf.

Eine wichtige Frage im Hinblick auf diese Darstellungsweise bezieht sich auf das Kombinationsspektrum des Typs. Dies erkennt man anhand des Auftretens fremder Episode an den Peripherien des Graphen. Im gegebenen Fall handelt es sich um Episoden aus dem Typ a650A - Der starke Hans.

Treten ein oder mehrere fremde Episoden an den Peripherien der ausgewerteten Episodenkette oder ihre Bruchteile auf, so entsteht die Notwendigkeit der gesamte Graph in die entsprechende Richtung zu ergänzen. Diese erfolgt durch das Eintragen des entsprechenden Präfix der gesuchten Episode ins Skript und das erneute Generieren des Graphen.[[14]](#footnote-14)

Dieses Mal handelt es sich um die Erweiterung des Graphen durch das Einfügen der Auswertungen der Episodengruppen Typ *a650A - Der starke Hans*. Der daraus generierte Graph sieht wie folgt aus:

Beim Typ a*650A - Der starke Hans* kann dieser gleich nach der ersten (*e301\_c\_verlust\_des\_OB*) oder/und zweiten (*e301\_e\_vom\_beginn\_der\_aktivitaet\_bis\_zum\_finden\_der\_raststaette\_oder\_eines\_eingangs*) Episode durch den Typ *a301 - Die drei geraubten Prinzessinnen* fortgesetzt werden. Es ist ebenfalls möglich, dass die Koppelung zwischen a301 und a650A erst dann stattfindet, wenn a650A als selbstständige Geschichte sich ziemlich weit entwickelt hat.

* *e650A\_c\_von\_der\_geburt\_des\_HD\_bis\_zum\_verlassen\_des\_hauses*

Der Held wird nach der Tier-Mensch bzw. Riesen-Mensch Beziehung geboren. Der Held befreit sich und seine Mutter aus der Gefangenschaft seines Vaters und macht sich auf den Weg zur Suche nach Abenteuer.

Somit kann der Typ a650A ohne weiteres an die Episode e301\_e anknüpfen und weiter als der Typ a301 fortgesetzt werden (siehe oben die Episode e301\_e\_).

* *e650A\_e\_verlassen\_des\_hauses\_bis\_zum\_treffen\_mit\_dem\_AN*

Im Zuge der Entwicklung des Typs als selbstständige Geschichte trifft der Held nicht die Männer mit außergewöhnlichen Eigenschaften, die ihn später als Weggefährten begleiten, sondern wird gezwungen das Haus wegen seines unstillbaren Appetites zu verlassen und nach dem Auftraggeber zu suchen, der ihm gegen seine Arbeit genug Essen geben kann.

* *e650A\_g\_besuch\_beim\_AN\_und\_reaktion\_des\_AN*

Der Auftraggeber lässt sich von Kraft des Helden beeindrucken und stellt ihn als seinen Knecht ein.

* *e650A\_i\_eine\_der\_vorletzten\_aufgaben*

Gleich nach der Abmachung hat der Held verschiedene Aufgaben zu erfüllen: viel Holz beschaffen, Zollerhebung vollstrecken usw.

* *e650A\_k\_die\_letzte\_aufgabe*

Das Erfüllen der Aufgaben mündet gewöhnlich in Ergebnisse, die den Auftraggeber in Schrecken und Angst versetzen. Die letzte Aufgabe die gezielt darauf ausgerichtet ist, dass der Held stirbt, endet mit einem Rückschlag für den Auftraggeber.

* *e650A\_m\_nach\_der\_erfuellung\_der\_letzte\_aufgabe\_bis\_zum\_ende\_der\_geschichte*

Nach der Auflösung der Abmachung mit dem Auftraggeber, kann die Geschichte ihr logisches Ende finden, oder es kann auch passieren, dass der freie Held nach anderem Abenteuer sucht. Trifft letzteres ein, so ist es möglich, dass der Held aus dem Typ a650A die neue Funktion des Helden vom Typ a301 übernimmt, indem er gleich nach der Trennung vom Auftraggeber mit den neuen Weggefährten trifft und gleich nach dem Aufschlagen des Nachtquartiers den Anlass fürs Treffen mit dem außergewöhnlichen Männlein vorbereitet.

Als Hilfsmittel für die einfache und schnelle Wahrnehmung der durch die Grafik visualisierten Normalform des Typs bzw. des Typenclusters gilt die farbliche Differenzierung der Verbindungslinien bzw. Kanten. Die roten Linien weisen auf häufig anzutreffende Knoten hin, die blaue Linien bezeichnen die mittelhäufige Beziehungen zwischen den Knoten und die grüngefärbten Linien entsprechen seltenen Koppelungen.[[15]](#footnote-15)

Es ist möglich die Auswertung durch das Entfernen seltener und somit unwesentliche Knoten zu entlasten und, soweit möglich, ein klareres Bild der wichtigsten Koppelungen erhalten. So stellt sich z.B. die obige Visualisierung durch die Korrektur des Häufigkeitsindexes von 1 auf 3 wie folgt dar:

Nach diesem Bild ist eindeutig ein Zusammenhang zwischen den Typen *a301* und *a650A* zu erkennen. Dabei trifft *a650A* ausnahmslos als Anfangsgeschichte auf, die später zu keinem, für diesen Typ charakteristischen Ende geführt wird. In einem anderen Fall wird die Episode *e301\_e\_* auch mit einer gewissen Intensivität als Anfang der Geschichte hervorgehoben, wobei eine dritte Variante des Anfangs von *e301\_c\_* eher zurückhaltend bleibt.

Nach dem Graph können also folgende zwei Episoden als Anfang für den Typ a301 alternieren:

* *e301\_c\_verlust\_des\_OB*

Nach dieser Episode beginnt die Geschichte damit, dass drei Königstochter in die Gewalt eines Ungeheuers geraten. Der König kündigt an, sie demjenigen zur Frau zu geben, der sie zurückbringt.

* *e650A\_c\_von\_der\_geburt\_des\_HD\_bis\_zum\_verlassen\_des\_hauses*

Der Held ist aus einer Tier-Mensch bzw. Riesen-Mensch-Beziehung geboren. Der Held befreit sich selbst und seine Mutter aus der Gefangenschaft seines Vaters und macht sich auf den Weg zur Suche nach Abenteuer.

Wie das Graph zeigt, bleibt die sich danach entwickelnde Geschichte stabil und kommt als der Typ a301 zum Ende (s. oben die Beschreibung der Episoden von e301\_e\_ bis e301\_o\_).

Damit ist eigentlich die Normalform für den Typ *a301 – Die drei geraubten Prinzessinnen* im deutschsprachigen Repertoire ergründet (natürlich unter dem Vorbehalt, dass die betrachteten Texte nur einen Bruchteil des Repertoires darstellen).

Nun führen wir dieselbe Analyse anhand des georgischen Repertoires durch. Das Graph für die Beziehungen aller Episoden ist hier allerdings so unübersichtlich, dass eine sinnvolle Visualisierung gleich mit der Häufigkeitsraten = 2 beginnen sollte.

Hier ist leicht zu erkennen, dass die Bindeglieder der Episodengruppe *e301* keinesfalls als beginnende oder abschließende Teile der gesamten Geschichte fungieren. Sie sind dementsprechend von Episoden anderer Typen umschlossen, und zwar von Episoden der Typen *a516* und *a300*. Tragen wir also entsprechende Präfixe an die passende Stelle des Skripts ein und lassen wir das Programm die Beziehungen erneut visualisieren.

Die Bruchteile der Kette werden zusammengebunden und wir sehen weitere fremde Episode an den Peripherien der erweiterten Kette. Das sind *eCoM*, *e303* und *e650A.* Nun übertragen wir auch diese Episodenpräfixe ins Skript und generieren eine erweiterte Episodenkette:

Durch das Bild wird klar, dass der Typ *a301* im georgischen Repertoire in engster Affinität mit den Typen *a516* und *a300* steht und zusammen mit ihnen ein gemeinsames Cluster bildet. Dabei kann das Cluster unterschiedlich ausgedehnt werden. Als typisch gilt der Ausfall der Episoden von *e516\_m* bis *e516\_w* am Ende der Geschichte. Ebenfalls sind weitere Veränderungen zu beobachten, z.B. die aus dem deutschsprachigen Repertoire uns gut bekannte Kontamination mit dem Typ *a650A*. Diese scheint allerdings im georgischen Repertoire eher auf einem Zufall zu beruhen und lässt sich bei weiterer Erhöhung der Häufigkeitsrate kaum noch beobachten.

Alles in Allem sieht die Realisierung des Typs a301 im georgischen Repertoire diametral anders aus als sein Pendant im westeuropäischen bzw. deutschen Märchenrepertoire.

Äußerst auffällig ist das Motiv des Frauenraubes. Dier passiert im Zusammenhang mit dem Typ *a516 - Der treue Johannes*. Hier beginnt die Geschichte mit der Episode der wundersamen Empfängnis (erster Knoten unten links). Als Handlungsabschnitt kommt die Gewinnung und dann der Verlust der Frau. Diesem folgt die Wiedergewinnung der Frau, dann die in der Unterwelt vollbrachten Heldentaten und schließlich der Flug zurück in die Oberwelt auf dem Rücken des Vogels, der dem Helden nach dem Ende des Fluges ein wunderbares Heilmittel schenkt. Der Helfer erfährt von den lauern den Gefahren und vernichtet mehrere kostbare Geschenke, die den Königssohn und seine Braut zum Tod führen können. Der Helfer erklärt dem Helden sein Verhalten und wird durch den Bruch des Schweigens zu Stein. Die Frau begießt den leblosen Leib des treuen Helfers mit dem Blut ihrer eigenen Kinder und holt ihn somit zurück ins Leben. Der Freund wiederbelebt die Kinder durch das Heilmittel wieder, das er vom Vogel zum Geschenk bekommen hatte (die ausführliche Beschreibung der Geschichte ist im Abschnitt 3.2 Markup zu finden).

Trotzt der schematischen Darstellung der somit erfassten Geschichten kann der dadurch erstellte Überblick über die Abfolge der Episoden sowie über die Kombination den Typen sehr dazu beitragen, die unterschiedliche Entfaltung der Typen übersichtlich zu halten. Das wiederum ermöglicht es uns, ohne größeren Aufwand Schlussfolgerungen zu ziehen, auch wenn die ohnehin schon unüberschaubare Datenmenge permanent anwächst.

1. Würzbach, Natascha (1999): Motiv. Enzyklopädie des Märchens Sp.946-954. [↑](#footnote-ref-1)
2. Als gutes Beispiel dafür gilt folgende Kontamination von Typen *ATU 425A - Das Ungeheuer als Bräutigam* + *ATU 563 - Tischlein deck dich*. [↑](#footnote-ref-2)
3. Die vor dem Beginn der Tätigkeit des Helden geborenen Geschwistern sind zuerst als falsche Rückeroberungsobjekte zu bestimmen. Erst nach ihrer Befreiung vom jüngeren Bruder begehen sie Missetat gegenüber ihn und übernehmen somit die Funktionen des Falschen Helden, bleiben wie vorher als fRE. [↑](#footnote-ref-3)
4. Der Terminus gilt für beliebige Lösungsbedürftige Zustand im Märchen: Schädigung, Mangelsituation, Druck etc. [↑](#footnote-ref-4)
5. vollständige Liste der Elemente s. GitHub [↑](#footnote-ref-5)
6. Wasser des Lebens, KHM 97; vgl. Blécourt, Willem de(2014): Wasser des Lebens. Enzyklopädie des Märchens. Sp. 509-514. [↑](#footnote-ref-6)
7. Gemeint werden geklonte und abgeleitete Wortlaute, die die *c* Attributwerte ausmachen (siehe den Abschnitt *Attribut c*) [↑](#footnote-ref-7)
8. Anhang 6.2.1 Künstlicher Assistent für die Annotation des Märchens [↑](#footnote-ref-8)
9. Anhang 6.2.1 Künstlicher Assistent für die Annotation des Märchens, Zellen #14-1. [↑](#footnote-ref-9)
10. W.o. Zellen # 16 – # 18. Als vorbildlich gilt die Anwendung von Needham, Mark 2016. [↑](#footnote-ref-10)
11. Repertoire entspricht den in der elektronischen (CD) Ausgabe der „Europäische Märchen“ zusammengetragene Texte unter dem Typnummer ATU 551. [↑](#footnote-ref-11)
12. Stellt ein Mitglied aus den gesuchten Motiven den Anfang oder das Ende eines Textes dar, so wird dieses durch extra Zeichen „Anf-“ und „End-“ gekennzeichnet. [↑](#footnote-ref-12)
13. Anhang 6.2.2 Künstlicher Assistent für die Auswertung der annotierten Daten, Zelle #3, Zeile 1. [↑](#footnote-ref-13)
14. w.o. Zeile 2. [↑](#footnote-ref-14)
15. Den momentan vorhandenen Konventionen zufolge sind die oben erwähnten Gradationen zwischen „selten“ und „häufig“ wie folgt festgehalten: grün = selten (<=33%), blau = mittel (>33% & <50%), rot = häufig (>= 50%). [↑](#footnote-ref-15)