

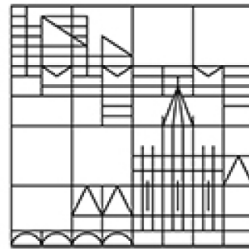
Eine computerlinguistische Untersuchung des Genitivschwundes

MASTERARBEIT

vorgelegt von
Christin Schätzle

an der

Universität
Konstanz



Geisteswissenschaftliche Sektion
Fachbereich Sprachwissenschaft
Hauptfach Speech and Language Processing

1. Gutachterin: Professorin Dr. Miriam Butt
2. Gutachter: Professor Dr. Frans Plank

Konstanz, im März 2013

Danksagung

Ich bedanke mich bei Professorin Dr. Miriam Butt für die intensive Betreuung meiner Masterarbeit und die tatkräftige Unterstützung bei der Themenfindung. Ein weiterer Dank geht an Professor Dr. Frans Plank für den gedanklichen Anstoß zur besseren Gestaltung meines Programmes und zur Interpretation der vorliegenden Ergebnisse.

Insbesondere möchte ich meiner ehemaligen Kommilitonin Joanna Müller danken, die sehr viel Energie in die Korrektur meiner Masterarbeit gesteckt hat und deren Verbesserungsvorschläge diese um einiges weiterentwickelt haben. Ebenso bedanke ich mich bei Inge Beck, die durch ihre Bemühungen zur Verbesserung der vorliegenden Arbeit beigetragen hat.

Inhaltsverzeichnis

1. Einleitung	1
2. Theoretischer Hintergrund.....	3
2.1 Bisherige Erkenntnisse zum Genitivschwund.....	3
2.1.1 Rückgang des adverbalen Genitivs	3
2.1.2 Rückgang des adnominalen Genitivs	5
2.1.3 Ursachen des Genitivschwundes.....	7
2.2 Die literarische Gattung Novelle.....	10
2.2.1 „Novelle“ von Johann Wolfgang von Goethe	11
2.2.2 „Der Tod in Venedig“ von Thomas Mann.....	12
2.2.3 „Die Entdeckung der Currywurst“ von Uwe Timm	13
3. Computerlinguistischer Hintergrund.....	15
3.1 Die Programmiersprache Perl.....	15
3.2 Part-of-speech Tagging	16
4. Durchführung	18
4.1 Struktureller Ablauf	18
4.1.1 Vorbereitung der Texte für das POS Tagging	18
4.1.2 Anwendung des TreeTaggers	21
4.1.3 Perlskript zur Untersuchung des Genitivschwundes im Deutschen.....	21
4.2 Probleme bei der Erstellung des Programmes	36
5. Ergebnisse der computerlinguistischen Analyse.....	39
5.1 Resultate der Genitivuntersuchung	39
5.2 Auswertung der erhobenen Daten	43
6. Diskussion.....	55
7. Fazit	66
Literatur- und Quellenverzeichnis	V
Anhang.....	VII

Abbildungsverzeichnis

Abbildung 1: Beispielausgabe des <i>TreeTaggers</i> (TreeTagger, 2013)	17
Abbildung 2: Programmaufruf und Dateiabfrage von <i>preprocess.pl</i> in der Betriebssystem-Shell.....	19
Abbildung 3: Programmausschnitt aus <i>preprocess.pl</i>	20
Abbildung 4: Dateikontrollstruktur aus <i>genitivex.pl</i>	22
Abbildung 5: Ermittlung der Anzahl aller in <i>\$datei1</i> enthaltenen Wörter.....	23
Abbildung 6: Erstellung der aus Wörtern und <i>POS Tags</i> bestehenden Bigramme.....	24
Abbildung 7: Programmaufruf und Dateiabfrage von <i>genitivex.pl</i> in der Betriebssystem-Shell	25
Abbildung 8: Erstellung größerer N-Gramme aus Bigrammen.....	26
Abbildung 9: Suchmuster für einen adnominalen Genitiv	27

Tabellenverzeichnis

Tabelle 1: Beispiele für verschiedene Genitivtypen	6
Tabelle 2: Beispiele für verschiedene Untergruppen des adnominalen Genitivs.....	7
Tabelle 3: Adnominale Genitive und deren prozentualer Anteil bezogen auf die gesamte Wortanzahl.....	43
Tabelle 4: Adnominale Genitive und deren prozentuale Anteile bezogen auf alle Genitive	44
Tabelle 5: Alle Genitivtypen und deren prozentuale Anteile bezogen auf die gesamte Wortanzahl.....	45
Tabelle 6: Prozentuale Anteile aller Genitivtypen bezogen auf alle Genitive.....	47
Tabelle 7: Unterschiedliche Genitivarten und ihre Konkurrenz	48

1. Einleitung

Der Genitivschwund stellt ein Problem dar, das in der deutschen Sprachwissenschaft heftig diskutiert wird. Es scheint, als verliere vor allem die Jugend die Fähigkeit, den Genitiv korrekt zu benutzen und wende sich lieber anderen Konstruktionen, wie dem possessiven Dativ, zu. Viele interpretieren dies als Verkümmern der deutschen Sprache. Es handelt sich jedoch nicht um ein Phänomen, das erst in jüngster Vergangenheit auftrat, denn unter anderem Donhauser (1998) und Fleischer & Schallert (2011) belegen bereits für ältere Sprachstufen einen Rückgang in der Genitivverwendung.

Das im Rahmen dieser Arbeit erstellte Perlskript soll dazu dienen, den Genitivschwund zu untersuchen und dabei auf Mittel der maschinellen Textverarbeitung zurückgreifen. Weiterhin gilt es, den Funktionsverlust des Genitivs innerhalb der letzten beiden Jahrhunderte zu analysieren und sein Fortschreiten zu belegen. Dazu wurden drei Novellen, die als Textsorte nahe an der Alltagssprache liegen und in eben genanntem Zeitrahmen entstanden sind, untersucht.

Als Beispiel der standarddeutschen Sprache oder vielmehr eines vorbildlichen Sprachgebrauchs wird „Novelle“ von Johann Wolfgang von Goethe, welche erstmals 1828 erschien, herangezogen. „Die Entdeckung der Currywurst“, eine Novelle von Uwe Timm, die aus dem Jahr 1993¹ stammt, dient als Referenz für die moderne deutsche Sprache. Zeitlich dazwischen liegt die dritte Novelle unter dem Titel „Der Tod in Venedig“ von Thomas Mann aus dem Jahr 1912. Geht der Genitiv innerhalb dieser Zeitspanne weiter zurück, so sollte die Untersuchung durch das in Perl erstellte Programm ergeben, dass der Genitiv in seiner Verwendung von Goethe über Mann bis hin zu Timm immer weiter sinkt. Um eine Verkümmern der deutschen Sprache in jüngster Vergangenheit zu demonstrieren, müsste wiederum der Rückgang von Goethe zu Timm deutlich bedeutender sein als jener von Goethe zu Mann. Zudem sollte erneut ein starker Verlust des Genitivs von Mann bis zu Timm aufgezeigt werden können.

Das Perlskript, welches zu diesem Zweck erstellt wurde, muss in der Lage sein, alle Genitive in den Texten zu identifizieren, auszugeben und zu zählen. Des Weiteren sollte das Programm auch solche Konstruktionen wahrnehmen und zählen, die anstelle eines Genitivs verwendet werden, damit deren Einfluss auf den Genitivschwund

¹ Es wird die 13. Auflage von 2008 herangezogen.

gemessen werden kann. Insgesamt sollten prozentuale Häufigkeiten errechnet werden, um textübergreifende Aussagen hinsichtlich des Genitivschwundes machen zu können.

Während der Erstellung des Perlskripts zur computerlinguistischen Untersuchung des Genitivschwundes wurde folgender Vorschlag Fleischer & Schallerts (2011) berücksichtigt:

„Um einer Lösung des Genitivproblems näher zu kommen, könnte es sich vielleicht als sinnvoll erweisen, möglichst viele Typen des Genitivs einheitlich zu erfassen.“

(Fleischer & Schallert, 2011: 100)

Somit sollen möglichst viele unterschiedliche Typen mit Hilfe des Programmes untersucht werden, um Aussagen darüber machen zu können, welche Genitivtypen besonders betroffen sind. Darüber hinaus könnten dadurch Ursachen für den Rückgang der einzelnen Arten festgestellt werden.

Damit die Untersuchung und deren Ablauf besser nachvollzogen werden können, werden zunächst der theoretische sowie der computerlinguistische Hintergrund dargestellt. Dazu wird im folgenden Kapitel als erstes genauer auf den Genitiv und dessen Schwund eingegangen. Anschließend wird der Novellenbegriff verdeutlicht, gefolgt von Beschreibungen der drei Novellen und deren Besonderheiten. Das Kapitel über den computerlinguistischen Hintergrund erläutert die Programmiersprache Perl und das zur Textbearbeitung genutzte *Part-of-Speech Tagging* näher. Ein weiteres Kapitel widmet sich der Durchführung, wobei die Erstellung und Verwendung des Programmes illustriert werden. Hierbei werden der strukturelle Ablauf, der die Vorbereitung der Texte, die Anwendung des *Part-of-Speech Taggings* und das Perlskript selbst beschreibt und anschließend die Probleme, die bei all diesen Schritten entstanden, verdeutlicht. Im Kapitel zu den Ergebnissen der computerlinguistischen Analyse werden die Resultate, die das Programm in eine Datei ausgibt, aufgeführt und ausgewertet. Diese werden dann im nächsten Kapitel diskutiert. Den Abschluss der Arbeit bildet das Fazit, in dem auf die anfangs gesetzten Ziele eingegangen wird und welches zusammenfassend berichtet, inwiefern diese erreicht wurden.

2. Theoretischer Hintergrund

Der Genitivschwund wurde in der Vergangenheit bereits eingehend untersucht. Dieses Kapitel dient zur Darstellung der bisherigen theoretischen Erkenntnisse und Thesen, die hinsichtlich des Genitivschwundes vorliegen. Zudem wird die literarische Gattung der Novelle vorgestellt und die drei Novellen, die zur Erstellung des Perlskripts herangezogen wurden, werden näher beschrieben.

2.1 Bisherige Erkenntnisse zum Genitivschwund

Fleischer & Schallert (2011) beschäftigen sich in Kapitel 6 ihres Buches „Historische Syntax des Deutschen: Eine Einführung“ detailliert mit dem Schicksal des Genitivs. Ebenso wie Karin Donhauser (1998) legen sie zunächst unterschiedliche Genitivtypen dar und analysieren anschließend deren Rückgang. Die Ausführungen in diesem Abschnitt über den Genitivschwund sind an deren Darstellungen angelehnt. Dabei werden der adverbale, sowie der adnominaler Genitiv und mögliche Gründe für deren Funktionsverluste näher behandelt.

2.1.1 Rückgang des adverbalen Genitivs

Die Verwendung des Genitivs als Objekt ist in der neuhochdeutschen Standardsprache nur noch äußerst selten aufzufinden. Dieser sogenannte adverbale Genitiv gilt heutzutage als unüblich. Diejenigen Verben, die einen Genitiv als Objekt zu sich nehmen können, werden als einer gehobenen Stilebene zugehörend angesehen, deren Klang archaisch ist (siehe Fleischer & Schallert, 2011).

Darüber hinaus gibt es nach Lenz (1996) nur noch 56 neuhochdeutsche Verben, die überhaupt einen Genitiv als Objekt haben können. Diese sind beispielsweise Verben wie „*sich einer Sache erdreisten*“ oder „*sich einer Sache bemächtigen*“ (Fleischer & Schallert, 2011: 83). Nach Donhauser (1998) existieren noch 40 neuhochdeutsche Verben, die als einziges Objekt ein Genitivobjekt zu sich nehmen. Betrachtet man die Untersuchungen von Rausch (1897) und Donhauser (1998) so ging diese Zahl im Laufe der Zeit stark zurück. Zwar war der Rückgang von 300 althochdeutschen genitivfähigen Verben bei Rausch (1897) beziehungsweise von 290 bei Donhauser (1998) zu 260 genitivfähigen Verben fürs Mittelhochdeutsche, die Rausch (1897) feststellte, nur gering, jedoch ist der Schwund bis hin zum Neuhochdeutschen beachtlich. Ausschlaggebend dafür waren vermutlich Entwicklungen im 14./15. Jahrhundert (siehe Fleischer & Schallert, 2011).

Auch Zählungen des Genitivobjekts, wie sie Lindgren (1969) vornahm, machen Aussagen über den Verlust des adverbalen Genitivs. Dabei wurde in verschiedenen Textauszügen bestehend aus 500 Satzgliedern zur gesamten deutschen Sprachgeschichte

gezählt, wie oft auf 100 Satzglieder ein Genitivobjekt erscheint. Lindgren (1969) fand hierbei heraus, dass bei Otfrid die relative Frequenz des Genitivobjekts 9,4 %, im Nibelungenlied 9,0 %, bei Luther 3,6 %, bei Goethe 1,2 % und bei Musil 1,0 % betrug, während in der Zeitschrift *Zeit* die Frequenz nur noch bei 0,2 % lag (siehe Fleischer & Schallert, 2011).

Diese Beobachtung und auch die Resultate Sommerfeldts (1966), der feststellt, dass zwar in Zeitungstexten aus dem Jahr 1860 das Genitivobjekt noch mit einer Frequenz von 1,5 % auftritt, im Jahr 1960 das Genitivobjekt aber lediglich noch in einer Frequenz von 0,6 % vorkommt, lassen darauf schließen, dass auch in jüngeren neuhochdeutschen Sprachstufen das Genitivobjekt mehr und mehr verschwindet.

Viele neuhochdeutsche genitivfähige Verben haben ferner die Eigenschaft, dass sie mehrere Objekttypen zu sich nehmen können, die somit in Konkurrenz zueinander stehen. So kann zum Beispiel bei einigen Verben ein Akkusativobjekt anstatt eines Genitivobjekts auftreten (Fleischer & Schallert, 2011: 83):

„ich entbehre seines Rates/ seinen Rat“

Zudem können auch Präpositionalobjekte, vor allem bei reflexiven Verben, einen Genitiv ersetzen (Fleischer & Schallert, 2011: 84):

„ich erinnere mich seiner/ ich erinnere mich an ihn“

Weiterhin kommt es vor, dass der adverbale Genitiv bei manchen Verben nur noch in feststehenden Wendungen auftritt (Fleischer & Schallert, 2011: 84):

*„er freut sich des Lebens vs. *er freut sich des Geschenkes“*

Außerdem ist nach Fleischer & Schallert (2011) in den Dialekten der neuhochdeutschen Sprache bis auf wenige archaische Ausnahmen ein vollständiger Abbau des adverbale Genitivs bis hin zu einem Drei-Kasus-System zu erkennen.

2.1.2 Rückgang des adnominalen Genitivs

Der adnominale Genitiv, oder auch das sogenannte Genitivattribut (siehe Fleischer & Schallert, 2011), ist in der neuhochdeutschen Standardsprache, im Gegensatz zum adverbale Genitiv, noch durchaus stabil und weit verbreitet. Deshalb sprechen Fleischer & Schallert (2011) davon, dass der Genitiv im Wesentlichen ein adnominaler Kasus sei. Nichtsdestotrotz geschah in den meisten deutschen Dialekten ein vollständiger Abbau des Genitivs und der adnominale Genitiv wird vor allem durch die *von*-Periphrase sowie den possessiven Dativ ersetzt. So wird beispielsweise die standarddeutsche Nominalphrase *das Spielzeug des Kindes* im Großteil der Mundarten durch die *von*-Periphrase *das Spielzeug vom Kind* beziehungsweise den possessiven Dativ *dem Kind sein Spielzeug* ersetzt. Dabei beschränkt sich der possessive Dativ laut Fleischer & Schallert (2011) nur auf belebte Nominalphrasen.

Kiefer (1910), der vorrangig Texte untersuchte, die nahe an der gesprochenen Sprache lagen, geht davon aus, dass Belege für die *von*-Periphrase bereits im 12. Jahrhundert zu finden sind. Jedoch ist es häufig möglich, dass *von* wohl eher wegen seiner spezifischen Semantik ausgewählt wurde und somit nicht als eindeutige Ersetzung eines adnominalen Genitivs gewertet werden kann (siehe Fleischer & Schallert, 2011).

Weiterhin ist denkbar, dass die *von*-Periphrase durch Reanalyse einer ehemals vom Verb abhängigen Präpositionalphrase zu einer zum vorangehenden Substantiv gehörenden Präpositionalphrase entstand. Als Beispielkontext hierfür wird bei Fleischer & Schallert (2011: 95) folgende Phrase angeführt:

„*da erging ein Gebot von Kaiser Augustus*“

Wahrscheinlicher ist hier zwar die Analyse, dass die Präpositionalphrase *von Kaiser Augustus* vom Verb abhängt, jedoch besteht die Möglichkeit, die Präpositionalphrase in attributiver Position zum vorhergehenden Substantiv zu interpretieren. Für die Reanalyse geben Fleischer & Schallert (2011: 95) folgendes Schema an:

„ältere Analyse: *da erging [ein Gebot][von Kaiser Augustus]*

jüngere Analyse: *da erging [ein Gebot [von Kaiser Augustus]]“*

Wird ein ambiger Kontext nach der Reanalyse auf andere Kontexte, die nur noch eine der beiden Interpretationen (in diesem Fall letztere) zulassen, übertragen, so spricht man von Extension. Für ältere Sprachstufen lässt sich aber kein eindeutiger Beleg finden, der zeigt, dass nach der Reanalyse einer *von*-Periphrase auch die Extension auf andere Kontexte stattfand. Diese Extension kann bewiesen werden, wenn nur noch

die Gesamtkonstituente innerhalb vom Satz verschoben werden kann, die ehemaligen Bestandteile aber nicht aus dem Gesamtkonstrukt entfernt werden können.

Auch der possessive Dativ, der den Genitiv ersetzen kann, kam vermutlich durch Reanalyse zustande (siehe Fleischer & Schallert, 2011). In diesem Fall können eine dativische Nominalphrase und eine darauffolgende, durch ein Possessivpronomen eingeleitete Nominalphrase in manchen Kontexten als zwei einzelne Nominalphrasen, die zum Verb gehören oder als eine einzige, zusammengehörende Nominalphrase analysiert werden. Die Extension dieser Konstruktion kann wieder durch Verschieben bewiesen werden oder auch dadurch, dass die Konstruktion in Zusammenhang mit einem Verb auftritt, welches eigentlich kein Dativobjekt zu sich nehmen kann. Mögliche Fälle des possessiven Dativs wurden zwar bereits in althochdeutschen Texten gefunden, es konnte aber nicht eindeutig nachgewiesen werden, dass es sich dabei um Formen des heutigen possessiven Dativs handelt. Laut Kiefer (1910) können erste eindeutige Belege für den possessiven Dativ erst im 12. Jahrhundert gefunden werden (siehe Fleischer & Schallert, 2011).

Neben dem adnominalen und dem adverbalen Genitiv kann der Genitiv im Neuhochdeutschen auch oder vielmehr immer noch in Abhängigkeit anderer Satzkonstituenten, beispielsweise einer Prä- beziehungsweise Postposition oder eines Adjektivs auftreten, wie in „*anlässlich des Treffens*“ oder „*seiner überdrüssig*“ (Fleischer & Schallert, 2011: 84). Der Rückgang dieser Genitivtypen soll in der vorliegenden Arbeit ebenfalls untersucht werden. Zusätzlich werden freie Genitive und Konjunktionen, Adjektive sowie Adverbien, die Genitiv enthalten, hinsichtlich des Genitivschwundes analysiert. All diese Genitivtypen sind in Tabelle 1 dargestellt und werden durch Beispiele näher erläutert.

Genitivtyp	Beispiel
Adnominaler Genitiv	<i>Der Wert des Wagens würde in den kommenden Jahren wenig zurückgehen.</i>
Adverbaler Genitiv	<i>Es bedurfte nur geringer Anstrengung, um den Kunden vom Kauf des Autos zu überzeugen.</i>
freier Genitiv	<i>Der Verkäufer stolzierte erhobenen Hauptes davon.</i>
Prä-/Postposition mit Genitiv	<i>Der Käufer machte sich wegen des hohen Treibstoffverbrauchs keine Sorgen.</i>
Adjektiv mit Genitiv	<i>Der Käufer war sich der enormen Geldsumme bewusst.</i>
Konjunktionen/Adjektive/Adverbien	<i>Keineswegs ließ er sich vom Kauf des Autos abbringen.</i>

Tabelle 1: Beispiele für verschiedene Genitivtypen

Des Weiteren wird der adnominale Genitiv in der folgenden Untersuchung in mehrere Untergruppen eingeteilt, die in Tabelle 2 durch Beispiele illustriert werden.

Untergruppe des adnominalen Genitivs	Beispiel
Adnominale Genitive mit definitivem Artikel	<i>das Spielzeug <u>des Kindes</u></i>
Adnominale Genitive mit indefinitem Artikel	<i>das Spielzeug <u>eines Kindes</u></i>
Adnominale Genitive mit Possessivpronomen	<i>das Spielzeug <u>seines Kindes</u></i>
Adnominale Genitive mit Indefinitpronomen	<i>das Spielzeug <u>irgendeines Kindes</u></i>
Adnominale Genitive mit Demonstrativpronomen	<i>das Spielzeug <u>dieses Kindes</u></i>
Adnominale Genitive mit Adjektiv ohne Artikel	<i>die Spielzeuge <u>glücklicher Kinder</u></i>
Vorangestellte adnominale Genitive	<i><u>des Kindes</u> Spielzeug</i>
Adnominale s-Genitive	<i><u>Lauras</u> Spielzeug</i>
Koordinierte adnominale Genitive	<i>das Spielzeug <u>der jüngsten Schwester und des ältesten Bruders</u></i>
Relativpronomen im Genitiv	<i><u>dessen</u> <u>deren</u></i>

Tabelle 2: Beispiele für verschiedene Untergruppen des adnominalen Genitivs

Anhand der Unterteilung in diese Untergruppen sowie mit der gesamten Unterteilung des Genitivs in bestimmte Typen soll der Genitivschwund detaillierter betrachtet werden können. Dabei könnte sich herausstellen, welche Genitivtypen maßgeblich von dem Funktionsverlust des Genitivs betroffen sind und hauptsächlich in ihrer Verwendung zurückgehen. Andererseits könnte sich zeigen, welche Genitivarten weniger von einem Rückgang erfasst werden und in ihrer Verwendung über den untersuchten Zeitraum hinweg stabil bleiben.

2.1.3 Ursachen des Genitivschwundes

Der Abbau des adverbale Genitivs kann im Wesentlichen auf zwei Ursachen reduziert werden. Zum einen schwindet der Genitiv, wenn Verben mit Genitivreaktion insgesamt aussterben. Zum anderen ist es möglich, dass, wie zuvor bereits erwähnt, andere Objekttypen, zum Beispiel Akkusativ-, Präpositional- oder Dativobjekte, den Genitiv ersetzen. Dabei kam es wahrscheinlich zunächst zur Variation mit den rivalisierenden Objekttypen, wodurch eine Konkurrenzsituation entstand. Hierfür führen Fleischer & Schallert (2011) interne als auch externe Faktoren auf.

Ein externer Faktor ist der Rückzug des adverbale Genitivs in bestimmte Stilschichten und Register. Extern deshalb, da der Einfluss außerhalb des Sprachsystems liegt. Der fehlende Genitiv in fast allen deutschen Dialekten lässt darauf schließen, dass der adverbale Genitiv zunächst in der Mundart und der Sprache des Volkes verloren ging, während er in der gehobenen und in der geschriebenen Sprache stabil blieb. Dann aber reduzierte er sich auch in der Schriftsprache auf gehobene Stilschichten und Register und der adverbale Genitiv wurde immer mehr zum Stilmittel (siehe Fleischer & Schallert, 2011).

Fischer (1987, 1992) untersuchte hierzu Leipziger Drucke von 1500 bis 1561 sowie Texte aus dem gesamten deutschen Sprachraum von 1570 bis 1630 und 1670 bis

1730. Er stellte fest, dass sich das Auftreten von Genitivobjekten vor allem in solchen Texten verringerte, die nahe an der gesprochenen Sprache geschrieben wurden und einen unterhaltenden oder privaten Charakter hatten. Damit drückt das Genitivobjekt als Stilmittel ein gehobeneres Register und eine höhere Stilschicht aus, was der Äußerung einen archaischen Klang verleiht. Darüber hinaus zeigte sich, dass in den letzten beiden analysierten Zeiträumen das Genitivobjekt auffallend häufig in Texten, die „sprachlich und stilistisch durch alte Textvorlagen, durch lateinische Ursprungstexte, durch Kanzleistil, Rechtssprache oder Sprache der Bibel in starkem Maße geprägt“ (Fischer, 1992: 333) sind, auftraten. Laut Fleischer & Schallert (2011) ist der adverbale Genitiv vor allem im Rechtsbereich bis heute noch stark vertreten.

Innerhalb des Sprachsystems spielt vor allem die Alternation mit anderen Objekttypen eine herausragende Rolle. Hierzu findet Donhauser (1998) bereits Belege in althochdeutscher Zeit. Die Autorin zeigt, dass die Unterteilung in zwei- und dreiwertige Verben einen Einfluss auf das Genitivobjekt hat. Bei den zweiwertigen Verben im Althochdeutschen gibt es nicht nur genitivfähige Verben, die als Objekt ausschließlich einen Genitiv nehmen, sondern auch solche, bei denen es zur Alternation zwischen einem Genitivobjekt und einem Akkusativ- oder Dativobjekt kommt. Donhauser (1998) identifiziert im Althochdeutschen mindestens 25 Verben, bei denen Genitiv und Akkusativ in Alternation stehen. Dreiwertige Verben können zwar zusammen mit einem Genitivobjekt und zusätzlich einem Dativ- oder Akkusativobjekt auftreten, jedoch gibt es hierbei nach Donhauser (1998) keine Objektalternation.

Diese Genitiv-Akkusativalternation, die bis in die mittelhochdeutsche Zeit reicht, bringt semantische Besonderheiten, unter anderem den Partitivitätseffekt, mit sich. Das Verb tritt dann mit Genitiv auf, wenn die Handlung, die es darstellt, nicht bis zum Schluss durchgeführt wurde und das Objekt nur teilweise von der Verbalhandlung betroffen ist. Hierbei spricht man vom partitiven Genitiv. Ist das Objekt aber vollständig davon betroffen, tritt das Verb mit Akkusativ auf. Im Neuhochdeutschen ist der partitive Genitiv nur selten zu finden, stattdessen wird der Betroffenheitsgrad des Objekts durch Anwendung beziehungsweise Streichung des bestimmten Artikels ausgedrückt (siehe Fleischer & Schallert, 2011). Zudem bekommen im Althochdeutschen Sätze, in denen Verben mit dieser Alternation einen Akkusativ als Objekt nehmen, eine resultative oder punktuelle Bedeutung. Die Handlung gilt somit als abgeschlossen. Bei nicht abgeschlossenen oder bei über einen längeren Zeitraum andauernden Handlungen wird Genitiv verwendet (siehe Donhauser, 1998). Fleischer & Schallert (2011) weisen darauf hin, dass in diesem Fall der Aspekt, der darüber Auskunft gibt, ob eine Verbal-

handlung abgeschlossen (terminativ) oder nicht abgeschlossen (nicht-terminativ oder durativ) ist, eine zentrale Rolle spielt.

Folglich schließt Donhauser (1998), dass aufgrund des Verlusts der Aspektkategorie während der mittelhochdeutschen Zeit auch Verben mit Objektalternation den Genitiv verlieren. Fleischer & Schallert (2011) fügen dem hinzu, dass der Rückzug in bestimmte Stilschichten hierdurch erst entstehen konnte und somit eher eine sekundäre Erscheinung ist. Der vorher eintretende Verlust der grammatischen Spezifität führt zum Verlust der genitivfähigen Verben. Allerdings kann diese Aspekthypothese den Rückgang des adverbalen Genitivs nicht vollständig erklären, da bei den dreiwertigen Verben keine vergleichbare Alternation gefunden wurde.

Der adnominale Genitiv hingegen ist lediglich in den Dialekten verschwunden, wobei er durch andere Konstruktionen ersetzt wurde. Wie bereits erwähnt, bleibt er jedoch in der standarddeutschen Sprache erhalten. Dennoch „kann insgesamt für die deutsche Sprachgeschichte von einem kontinuierlichen Verlust des Genitivs gesprochen werden“ (Fleischer & Schallert, 2011: 99).

Richtet man sich nach Behaghel (1923), so liegt die Ursache für den allgemeinen Rückgang des Genitivs in der sogenannten Nebensilbenschwächung. Dabei wurden alle vollen althochdeutschen Nebensilbenvokale im Mittelhochdeutschen zum Zentralvokal /ə/ (Schwa). Diese lautliche Veränderung zog morphologische Konsequenzen mit sich und der Genitiv fiel in manchen Paradigmen mit anderen Kasus zusammen. Jedoch kann der Genitivschwund damit nicht vollständig belegt werden, vor allem starke Maskulina und Neutra bilden bis heute den Genitiv Singular auf -(e)s und bieten keine Grundlage für die Nebensilbenschwächung als Ursache des Genitivschwundes (siehe Fleischer & Schallert, 2011). Auch Karin Donhauser (1998) gibt an, dass diese Schwächung kein ausreichender Beleg ist, da die Kasuszeichnung bis zur heutigen Zeit vom Artikel übernommen wird, der zumindest Genitiv Singular und Plural von Nominativ und Akkusativ unterscheidet.

Wie zu erkennen ist, konnten bis heute keine der Erklärungen für den Genitivschwund eindeutig belegt werden. Donhauser (1998) warnt zudem davor, das Genitivproblem bereits als gelöst zu betrachten. Somit sind an dieser Stelle weitere Untersuchungen nötig.

2.2 Die literarische Gattung Novelle

Hugo Aust (2012) gibt unterschiedliche Definitionen für den Novellenbegriff an. Eine dieser Definitionen, die ebenfalls im Duden zu finden ist und in der vorliegenden Arbeit übernommen wird, besagt, dass eine Novelle eine „Erzählung kürzeren oder mittleren Umfangs, die von einem einzelnen Ereignis handelt und deren geradliniger Handlungsablauf auf ein Ziel hinführt[, ist]“ (Novellendefinition Duden, 2013).

Prototypische Kennzeichen der Novelle sind das Gespräch, die Absicht des Gesprächs, das Erzählen und dessen Redegüte. Dabei geht jede Novelle aus einem Gespräch hervor oder besitzt zumindest den erzählerischen Charakter eines solchen. Somit erhält die Novelle nach Aust (2012) Merkmale der Geselligkeit und Mündlichkeit. Die Absicht besteht darin, wie aus der angeführten Novellendefinition hervorgeht, aus dem Gespräch heraus ein bestimmtes Ziel zu verfolgen und zu erreichen. Das Erzählen gibt hierbei durch seine gestalterische Kraft die Möglichkeit „Geschehenes zu erfassen, zu ordnen, der Reihe nach und folgerichtig wiederzugeben“ (Aust, 2012: 7).

Des Weiteren sind bei Aust (2012) folgende Kriterien zur Abgrenzung des Novellenbegriffs aufgeführt: Länge, Begebenheit, Konzentration, Rahmen, Erzählen nach Mustern und Sammelbarkeit. Gebräuchlicherweise gilt als Umfang einer Novelle der Umfang einer Erzählung mittlerer Länge. Novellen erzählen dabei von einer unerhörten beziehungsweise außerordentlichen, neuen, wahren und einzigartigen Begebenheit. Die Konzentration beinhaltet den Wendepunkt und das Symbol einer Novelle.

„Einheit und Fülle, die das Symbol – nach Goethes Auffassung – nicht nur anzeigt, sondern auch unmittelbar ist, geben der kürzeren Erzählform den nicht auslotbaren Sinn-Raum, verleihen ihrem engen Ausschnitt die unendliche Ausdehnung, ihrer Fixierung ans Einzelne und Einmalige die allgemeine Geltung.“

(Aust, 2012: 16)

Die aus dem Gespräch entstandene Erzählsituation bietet den Rahmen einer Novelle, der von außen auf das Innere der Erzählung wirkt. Die Einhaltung bestimmter Erzählmuster und die Kürze der Novelle bilden die Grundlage für deren Sammelbarkeit (siehe Aust, 2012).

In den folgenden Abschnitten werden die drei Novellen, auf denen das Perlskript zur Untersuchung des Genitivschwundes basiert, näher besprochen, um die Gründe für deren Auswahl zu verdeutlichen. Zusätzlich werden der typische Novellencharakter sowie bestimmte sprachliche Merkmale herausgearbeitet, damit im weiteren Verlauf der vorliegenden Arbeit darauf referiert werden kann.

2.2.1 „Novelle“ von Johann Wolfgang von Goethe

„»Wissen Sie was,« sagte Goethe, »wir wollen es die ›Novelle‹ nennen; denn was ist eine Novelle anders als eine sich ereignete, unerhörte Begebenheit. Dies ist der eigentliche Begriff, und so vieles, was in Deutschland unter dem Titel Novelle geht, ist gar keine Novelle, sondern bloß Erzählung oder was Sie sonst wollen.«“

(Eckermann, 1981: Kap. 79)

Auch Aust (2012) stimmt damit überein, dass der Titel von Goethes Werk durchaus passend gewählt ist. So spricht er vom Novellentitel als ein „Spiel mit verschiedenen literarischen Formen“ (Aust, 2012: 95). Die Charakteristik einer Novelle wird hier nicht durch einen Rahmen verdeutlicht, sondern die Identität der Novelle kommt eben im Aufbrechen dieses Rahmens als Metamorphose zum Vorschein (siehe Aust, 2012).

In Goethes „Novelle“ geht es zunächst um einen Fürsten, der zur Jagd gehen möchte und sich deshalb von seiner Gemahlin verabschiedet. Um die Fürstin bei Laune zu halten, schlägt er ihr vor, einen Spazierritt zu unternehmen. Diese nimmt den Vorschlag an und beginnt wenig später einen Ausritt zu den Ruinen einer alten Stammburg. Begleitet wird sie dabei von ihrem Stall- und Hofjunker Honorio und dem fürstlichen Oheim Friedrich. Die erste Station ihres Spazierritts soll der städtische Jahrmarkt sein. Beim Jahrmarkt angekommen treffen sie auf einen brüllenden Löwen und einen Tiger, die dort in einem Käfig zur Schau stehen. Anschließend setzen sie ihre Reise zur Stammburg fort. Dort eingetroffen entdeckt Honorio mit Hilfe eines Fernrohrs, dass der soeben besuchte Jahrmarkt in Flammen steht. Auf der Rückreise zur fürstlichen Residenz erscheint plötzlich ein Tiger vor ihnen, der durch den Brand vom Jahrmarkt fliehen konnte. Dieser scheint die Fürstin angreifen zu wollen, doch Honorio ist zur Stelle und tötet den Tiger mit einem Schuss in den Kopf.

Kurz darauf erscheinen die exotischen Besitzer des Tigers, vermutlich die Wärterin und deren Sohn, ein junger Knabe mit einer Flöte. Die Wärterin beklagt lautstark den unnötigen Tod ihres Tieres. Dabei wird auch der Besonderheit der fremden Sprache dieser Frau Beachtung geschenkt:

„Eine natürliche Sprache, kurz und abgebrochen, machte sich eindringlich und rührend. Vergebens würde man sie in unsern Mundarten übersetzen wollen; den ungefähren Inhalt dürfen wir nicht verfehlen [...].“

(Goethe, 1828: Kap. 5)

Nun erscheint der Fürst mit seinem Jagdgefolge. Zudem tritt der Mann der Wärterin auf und verkündet, dass auch der Löwe aus seinem Käfig entkommen sei. Wenig später wird klar, dass sich der Löwe in der soeben besuchten Burgruine friedlich sonnt. Aus Angst auch seinen Löwen zu verlieren, bittet der Wärter den Fürsten darum, ihn auf seine eigene Weise wieder einfangen zu dürfen. So soll der Sohn den Löwen durch

sein Flötenspiel besänftigen, damit der Vater den Löwen anschließend in seinen Käfig locken kann. Erstaunlicherweise erlaubt der Fürst diese sonderbare Methode.

Bald gelingt es dem Jungen und seinem Vater den Löwen durch Flötenspiel und Gesang zu besänftigen und langsam aus der Stammburg hinaus zu locken. Singend zieht der Junge, gefolgt von dem scheinbar zahmen Löwen, weiter. Die Novelle endet mit einer Zeile aus dem Lied des Knabens:

„und so geht mit guten Kindern selger Engel gern zu Rat, böses Wollen zu verhindern,
zu befördern schöne Tat. So beschwören, fest zu bannen liebem Sohn ans zarte Knie
ihn, des Waldes Hochtyrannen, frommer Sinn und Melodie“.

(Goethe, 1828: Kap.7)

Dieser hochlyrische Schluss ist nach Goethes Ansicht nötig, „denn nach der pathetischen Rede des Mannes, die schon poetische Prosa ist, mußte eine Steigerung kommen, ich mußte zur lyrischen Poesie, ja zum Liede selbst übergehen“ (Eckermann, 1981: Kap.77).

2.2.2 „Der Tod in Venedig“ von Thomas Mann

„Der Tod in Venedig“ von Thomas Mann scheint zwar gattungsgerecht keine mustergültige Novelle zu sein, jedoch hat der Autor diese und einige weitere Erzählungen bewusst, wohl aufgrund deren Objektivität, als solche bezeichnet. „Der Tod in Venedig“ ist dabei diejenige Erzählung, die (zusammen mit einer weiteren) einer Novelle am meisten entsprach (siehe Aust, 2012). Hermann Kurzke (2009) weist auf die Tatsache hin, dass noch heute Unterrichtsbehörden ihren Schülern den Stil Thomas Manns näherbringen, da vor allem „Der Tod in Venedig“ in einem perfekten klassizistischen Stil geschrieben ist.

Die Hauptfigur der Novelle ist Gustav von Aschenbach, ein verwitweter und alternder Schriftsteller, der ein hohes Ansehen genießt und aufgrund seiner herausragenden Leistungen geadelt wurde. Es wird davon ausgegangen, dass die Figur des Aschenbachs autobiographische Züge Thomas Manns trägt (siehe Kurzke, 2009). Den Beschluss auf Reisen zu gehen trifft Aschenbach, nachdem er während eines Spaziergangs auf einen Wanderer aufmerksam wird, der das Reisefieber in ihm weckt.

Zunächst führt ihn seine Reise auf eine Insel der Adria, da er hofft, dort sein Verlangen nach dem Fremdartigen und Bezuglosen zu stillen. Dies gelingt ihm jedoch nicht und der enttäuschende Strand, das schlechte Wetter und die österreichische Hotelgesellschaft führen dazu, dass er sich anderthalb Wochen später auf einem Dampfer nach Venedig befindet. In Venedig angekommen, wird Aschenbach in seinem Hotel beim Abendessen auf eine polnische Familie und vor allem auf deren schätzungsweise

vierzehnjährigen Sohn Tadzio, der durch seine Schönheit und Verletzlichkeit besticht, aufmerksam.

Die Tage darauf hat Aschenbach sehr unter dem schwülen Wetter und dem Scirocco zu leiden, was ihm zusätzlich das Gefühl gibt, alt zu sein. Lediglich die Momente, in denen er Tadzio beobachtet, spenden ihm Trost. Mehr und mehr richtet sich sein Tagesablauf nach jenem von Tadzio. Nichtsdestotrotz beschließt Aschenbach wegen seiner andauernden Schwindel- und Fieberanfälle sofort abzureisen. Da sein Koffer aber am Bahnhof in die falsche Richtung geschickt wurde, beginnt Aschenbach seine Abreise noch einmal zu überdenken und entscheidet sich, um Tadzios Willen in Venedig zu bleiben.

Nun beginnt Aschenbach sich seiner Liebe zu Tadzio hinzugeben und verfällt diesem völlig. Aschenbach spricht über Sokrates und Phaidros, womit Thomas Mann seinen literarischen Bezug zur Neuklassik verdeutlicht, indem der Autor auf Elemente der Antike referiert (siehe Kurzke, 2009). Auch die immer größer werdende Gefahr der Cholera, die sich nach und nach über Venedig ausbreitet, ignoriert Aschenbach.

Damit er Tadzio gefällt und auch um sich selbst eine jüngere Identität vorzugaukeln, lässt er sich seine ergrauten Haare färben und seine Falten wegschminken. Er verliert damit seine Selbstachtung und gibt sich ganz seiner bürgerlichen Herkunft hin. Er beginnt Tadzio durch die schmalen Gänge Venedigs zu verfolgen, wird dabei sogar von der Familie Tadzios ertappt und kauft zur Tarnung Erdbeeren am Straßenrand, die vermutlich mit der Cholera infiziert sind. Kurz darauf stirbt Aschenbach in seinem Liegestuhl, während er Tadzio am Strand beobachtet. Es scheint ihm, als ob der Junge ihm zuwinke und ihn dazu auffordere, ihm ins Meer zu folgen.

2.2.3 „Die Entdeckung der Currywurst“ von Uwe Timm

In „Die Entdeckung der Currywurst“ von Uwe Timm berichtet ein Ich-Erzähler, wie er von der Hauptfigur Frau Brücker die Geschichte zur Entstehung der Currywurst erfährt. Genaugenommen verfügt die Geschichte über zwei Erzähler, denn auch Frau Brücker nimmt eine erzählende Funktion ein. Hinzu kommt der formal mehrfach kodierte Text, „der genau jenen Übergang vom Alltagserzählen ins Kunstwerk, die Wendung vom Nahhorizont in den Fernhorizont, den Übergang von den »alltäglichen Dingen der Lebenswelt zum Monument« [...] sinnfällig macht“ (Hielscher, 2007: 149).

Die Novelle lässt zusätzlich einige autobiographische Rückschlüsse auf Uwe Timm zu. Die Charakterzüge der Hauptfigur Lena Brückers basieren auf denen der Tante und der Mutter Timms. Außerdem tritt die Figur des Vaters von Uwe Timm zu Zeiten des Schwarzmarkts in der Novelle auf (siehe Hielscher, 2007).

Der Ich-Erzähler, der einst in Hamburg lebte, besucht Frau Brücker im Altenheim, um zu erfahren, wie sie auf das Rezept der Currywurst gestoßen ist. Die inzwischen erblindete, alte Frau beginnt mit einer ausschweifenden Erzählung. Diese beginnt in den letzten Kriegstagen des Jahres 1945. Lena Brücker leitet zu dieser Zeit eine Kantine in Hamburg und trifft bei einem Kinobesuch auf den deutlich jüngeren Marinesoldaten Bremer, der sich anschließend bei ihr versteckt und fahnenflüchtig wird. Es entwickelt sich eine Liebesgeschichte zwischen Brücker und dem jungen Bremer. Nach Hielscher (2007) ist die unerhörte Begebenheit, die für die Novellenkennzeichnung verantwortlich ist, die Lügengeschichte, die Lena Brücker Bremer aufischt, indem sie ihm nach der Kapitulation Deutschlands das Ende des Krieges verschweigt und vorgibt, ihn weiter verstecken zu müssen. Hielscher (2007) bietet hier einen Vergleich zum Neuklassizismus, denn in der griechischen Antike wurde Kalypso sieben Jahre lang von Odysseus auf der Insel Ogygia festgehalten.

Nachdem Lena Brücker in einem Streit gesteht, dass der Krieg beendet ist, verschwindet Bremer. Hinzu kommt, dass sie auch in der Kantine nach dem Krieg nicht mehr gebraucht wird. Somit benötigt sie eine neue Möglichkeit, sich ihren Lebensunterhalt zu verdienen. Um sich eine Zukunft mit einem Imbissbudenstand aufzubauen, betreibt sie fleißig Tauschhandel mit den Engländern und erhält Würste, sowie Currypulver und Ketchup. Durch ein Missgeschick stolpert sie samt ihrer Errungenschaften die Treppe hinauf und die Ketchupflaschen zerbrechen. Als sie ihre von Ketchup und Curry verschmierten Finger ableckt, ist sie überwältigt von diesem besonderen Geschmack. Deshalb kann nicht von der Erfindung, sondern bestenfalls von der Entdeckung der Currywurst gesprochen werden:

„Dass die Currywurst zufällig entdeckt und nicht intentional erfunden wird, betont das Spontane, quasi Objektgeleitete, das von den Dingen Angesprochene und Reaktive, aus dem eine neue sinnliche Erfahrung und damit eine neue Weltsicht entspringt.“

(Hielscher, 2007: 152)

Als der Ich-Erzähler ein letztes Mal zu Frau Brücker fährt, wird ihm mitgeteilt, dass diese bereits verstorben ist. Sie hinterließ ihm jedoch ein Paket, welches den Pullover, den sie während seiner Besuche strickte, enthält und auch ein vergilbter Zettel ist darin zu finden. Auf der einen Seite deszettels stehen die Zutaten der Currywurst in Frau Brückers Handschrift. Auf der anderen Seite sind Teile eines Kreuzworträtsels zu sehen, dass Bremer zu lösen versuchte. Eines der Lösungsworte lautet Novelle.

3. Computerlinguistischer Hintergrund

Dieses Kapitel dient zur Darstellung der computerlinguistischen Details, die den Hintergrund für die vorliegende Untersuchung bilden. Hierzu wird an dieser Stelle näher auf die Programmiersprache Perl eingegangen. Außerdem wird das zur vorherigen Textbearbeitung genutzte *Part-of-Speech Tagging* erklärt.

3.1 Die Programmiersprache Perl

Die Programmiersprache Perl wurde Mitte der 80er Jahre des letzten Jahrhunderts vom Linguisten Larry Wall entwickelt. Perl eignet sich hervorragend für die Programmierung von Programmen, die innerhalb weniger Minuten erstellt werden sollen, bietet aber genauso eine gute Möglichkeit, lange und ausschweifende Programme zu erstellen, die mehrere Jahre in ihrer Entwicklung benötigen (siehe Schwartz et al., 2005).

Vorrangig dient Perl zur Textverarbeitung, ist aber auch in der Lage, andere Aufgaben zu erfüllen. Zwar ist Perl nicht einfach zu erlernen, dafür ist es aber einfach in seiner Benutzung. Mit seiner Aussage „Perl is easy to use, but sometimes hard to learn“ (Schwartz et al., 2005: 5) weist Larry Wall darauf hin, dass er sich bei der Konzipierung der Programmiersprache zu Gunsten des Programmierers und nicht des Schülers entschieden hat. Damit ist gemeint, dass man Perl nur einmal zu erlernen braucht, was langwierig und schwer sein kann, das Erlernte dafür aber danach immer und immer wieder anwenden kann. Die Programmiersprache bietet somit „mimum fuss and maximum enjoyment“ (Christiansen et al., 2012: 3).

Ein weiterer Vorteil von Perl ist die Schnelligkeit, denn die Entwickler lassen nur solche neuen Bestandteile zu, die eine kurze Arbeitszeit aufweisen. Perls gängiges Logo, welches ihr vom O'Reilly Media Verlag gegeben wurde und auch auf dem Cover des Buchs „Programming Perl“ (Christiansen et al, 2012) zu sehen ist, stellt das Kamel dar. Dieses Tier wurde gewählt, weil es zwar hässlich ist, allerdings sogar unter extremen Bedingungen hart arbeiten und alle gestellten Anforderungen erfüllen kann. Diese Eigenschaften können ebenfalls Perl zugeschrieben werden (siehe Schwartz et al., 2005).

Perl ist eine *high-level* Programmiersprache, denn der Programmiercode ist oftmals nur ein bis drei Viertel mal so lang wie der entsprechende Code in C, was eine Zeiteinsparung beim Schreiben, Lesen, Fehler beseitigen und in der Instandhaltung des Programmes bedeutet. Außerdem können auch Programme in Perl geschrieben werden, die für den Nutzer nicht lesbar sind (siehe Schwartz et al., 2005).

Auf den meisten Rechnern mit Unix-Betriebssystemen ist Perl bereits vorinstalliert. Doch auch Windows-Nutzer können Perl kostenlos herunterladen (siehe Perl, 2013). Das Perlskript zur Untersuchung des Genitivschwundes basiert auf *ActivePerl* in der Version 5.14.2. Da Perlprogramme im Wesentlichen Textdateien sind, erstellt und bearbeitet man sie am besten in einem Texteditor. Für die Erstellung des Perlskripts im Rahmen dieser Arbeit wurde der *Komodo Editor*, Version 7.1, benutzt, der bei *Active State* (siehe Komodo Editor, 2013) gratis zur Verfügung steht.

3.2 Part-of-speech Tagging

Bevor das eigentliche Programm auf die drei Novellen angewendet werden kann, müssen diese zunächst mit Hilfe des von Helmut Schmid im Jahre 1995 am Institut für Maschinelle Sprachverarbeitung der Universität Stuttgart entwickelten *TreeTagger* annotiert werden (siehe TreeTagger, 2013). Dabei unterlaufen die Texte dem Prozess des sogenannten *Part-of-speech Taggings*², wobei jedem Wort die zugehörige Wortart (*Part-of-speech*) zugeordnet wird:

„Part-of-speech tagging (or just tagging for short) is the process of assigning a part of speech or other syntactic class marker to each word in a corpus.”

(Jurafsky & Martin, 2009)

Bei der deutschen Version des *TreeTagger* entsprechen die *Part-of-speech Tags*³, also die Kürzel die zur Wortartenkennzeichnung benutzt werden, denen des Stuttgart-Tübingen-*Tagsets*⁴, welches unter STTS Table (2013) nachgeschlagen werden kann. Ausführliche Informationen zum *Tagset* und zur Wortartzuweisung sind außerdem unter STTS Guide (2013) zu finden. Dem *Tagging*-Prozess geht ein Tokenisierungsprozess voraus, bei dem der Text in einzelne Worte getrennt wird und ebenso die Satzzeichen zur nächst kleineren Einheit werden. Somit erhalten nicht nur die Wörter, sondern auch die Satzzeichen ein *Part-of-speech Tag*. Der *TreeTagger*, der auf der Homepage des Instituts für Maschinelle Sprachverarbeitung der Universität Stuttgart (siehe TreeTagger, 2013) kostenlos heruntergeladen werden kann, annotiert nun innerhalb des Textes nicht nur jedes Wort (beziehungsweise Satzzeichen) mit der entsprechenden Wortart, sondern gibt dazu noch Informationen über das jeweilige Lemma an. Somit steht in der Ausgabe des Programmes in jeder Zeile das Wort, dessen *POS Tag* und sein Lemma. Ein Beispiel für die Ausgabe des *TreeTaggers* wird in Abbildung 1 aufgeführt. Diese Ausgabe ist ein Beispiel für die Verwendung der englischen Version des *TreeTaggers*. Hierbei werden die *Part-of-speech Tags* des *Penn-Treebank-Tagsets* benutzt, die unter Penn Treebank (2013) genauer erklärt werden.

² kurz: *POS Tagging*

³ kurz: *POS Tags*

⁴ kurz: *STTS*

Sample output:

word	pos	lemma
The	DT	the
TreeTagger	NP	TreeTagger
is	VBZ	be
easy	JJ	easy
to	TO	to
use	VB	use
.	SENT	.

Abbildung 1: Beispielausgabe des *TreeTaggers* (TreeTagger, 2013)

Der *TreeTagger* ist für mehrere Sprachen erhältlich, indem verschiedene *Tagsets* den jeweiligen Sprachen entsprechend benutzt werden. Zudem kann der *TreeTagger* auch auf weitere Sprachen erweitert werden, solange ein entsprechendes Lexikon und ein zuvor manuell annotierter Trainingskorpus vorliegen (siehe TreeTagger, 2013).

4. Durchführung

Um den Genitivschwund umfassender untersuchen zu können, muss ein Programm in Perl geschrieben werden, das den nötigen Anforderungen entspricht. Zudem müssen die Texte für die Verarbeitung, unter anderem durch den eben erwähnten *TreeTagger*, vorbereitet werden. Der strukturelle Ablauf der Programmierung und auch die einzelnen Schritte, die zur Erstellung nötig waren, werden im folgenden Abschnitt verdeutlicht. Anschließend werden die Probleme, die während der Programmierung entstanden und die es zu beseitigen galt, angesprochen.

4.1 Struktureller Ablauf

Der erste Schritt innerhalb der Untersuchung besteht darin, die Novellen für das *POS Tagging* durch den *TreeTagger* vorzubereiten, um unbrauchbare Ergebnisse zu vermeiden. Daraufhin werden die Novellen entsprechend durch den *Tagger* annotiert, damit sie dann von dem tatsächlichen Hauptprogramm analysiert werden können. Der strukturelle Aufbau des Perlskripts zur Untersuchung des Genitivschwundes wird am Ende dieses Abschnitts dargelegt.

4.1.1 Vorbereitung der Texte für das POS Tagging

Zunächst werden die Novellen, die als Textdatei unter *goethe_novelle.txt* für „Novelle“ von Goethe, *mann_todinvendig.txt* für „Der Tod in Venedig“ von Mann und entsprechend *timm_currywurst.txt* für „Die Entdeckung der Currywurst“ von Timm ohne Vorwort und weitere nicht-inhaltsbezogene Angaben gespeichert sind, für den darauffolgenden Prozess des *Part-of-speech Taggings* vorbereitet. Diese Texte sind unter anderem auf der CD-ROM in Anhang 3 zu finden.

Zur Vermeidung von späteren Tagfehlern, wie etwa der Zuweisung der falschen Wortart, sollen die Dateien vor der Anwendung des *TreeTaggers* von bestimmten Zeichen bereinigt werden. Dafür wurde das Programm *preprocess.pl*⁵ erstellt, welches die Texte für die Zuweisung der *POS Tags* vorbereitet.

Der Programmaufruf kann unter anderem in der Betriebssystem-Shell erfolgen, nachdem man als erstes per Befehl *cd (change directory)* und der Pfadangabe des Ordners, in dem das Programm gespeichert ist, in eben diesen wechselt. Danach gibt man *perl* und den Namen der Programmdatei, in diesem Fall *preprocess.pl*, an und das Programm startet.

Zu Beginn wird der Benutzer mittels des *print-Statements* darüber informiert, wozu das Programm dient. Nach dem *Statement* steht in Anführungszeichen der auszugebende

⁵ ebenfalls in Anhang 3

String, der einer Zeichenkette entspricht (siehe Jurafsky & Martin, 2009). Der Ausdruck `\n` bezeichnet hierbei einen Zeilenumbruch⁶. Die Schrägstriche `\` vor den Anführungszeichen bewirken, dass diese nicht als Beginn oder Ende eines *Strings*, sondern als tatsächliche Anführungszeichen interpretiert werden.

Danach bietet das Programm dem User drei Auswahlmöglichkeiten in Bezug auf die Novellen. Die Auswahl wird per `<STDIN>` auf die skalare Variable `$dat1`, die einen Wert enthält, eingelesen. *STDIN* ist ein sogenanntes vordefiniertes *Filehandle* in Perl und führt in Verbindung mit dem *Line Reading* oder *Angle Operator* `<>` dazu, dass die vorherige Eingabe in der Kommandozeile als Input genommen wird (siehe Christiansen et al., 2012). Anschließend entfernt die Funktion *chomp* den durch die Eingabe entstandenen Zeilenumbruch vom Input.

Daraufhin wird unter Beibehaltung desselben Prinzips nach dem gewünschten Namen der Outputdatei gefragt. Die Eingabe wird wiederum aus der Kommandozeile per `<STDIN>` auf eine Variable, hier `$dat2`, eingelesen. Das Entfernen des Zeilenumbruchs erfolgt in diesem Fall ebenfalls per *chomp*. Abbildung 2 stellt den Programmaufruf sowie die gesamte Dateiabfrage für den Benutzer in der Betriebssystem-Shell inklusive Eingabe dar.

```
C:\Users\user>cd Genitivschwund
C:\Users\user\Genitivschwund>perl preprocess.pl
Dieses Programm dient zur Vorbereitung der Novellen "Novelle" von Johann
Wolfgang von Goethe, "Der Tod in Venedig" von Thomas Mann und "Die Entdeckung
der Currywurst" von Uwe Timm fuer die Bearbeitung durch den Iree Tagger.

Welche Datei moechten Sie bearbeiten? Zur Auswahl stehen goethe_novelle.txt,
mann_todinvenedig.txt und timm_currywurst.txt.
mann_todinvenedig.txt

Wie soll die bearbeitete Datei heissen? Bitte geben Sie den Dateinamen
im Format dateiname.txt ein.
venedig.txt

C:\Users\user\Genitivschwund>_
```

Abbildung 2: Programmaufruf und Dateiabfrage von *preprocess.pl* in der Betriebssystem-Shell

An diesem Punkt beginnt die tatsächliche Arbeit des Programmes. Diesen Abschnitt zeigt der Programmauszug in unkommentierter Form in Abbildung 3. Mit Hilfe von *open* wird die Inputdatei geöffnet. Als nächstes beginnt eine *while*-Schleife. Dabei wird der in geschweiften Klammern stehende Schleifenblock solange ausgeführt, wie die Bedingung in den Klammern nach dem *while-Statement* erfüllt ist. In diesem Abschnitt bedeutet das, dass während es noch Input von `$dat1` gibt, womit durch `<FILE1>` zugegriffen wird, der darunter stehende Block erfüllt wird.

⁶ In Abschnitt 4.1.3 folgt mehr zu regulären Ausdrücken.


```
open FILE1, "$dat1";

while (<FILE1>) {
    chop;
    tr/''''»«//d;
    tr/-/ /;
    tr/-/ /;
    foreach $w(split) {
        push @words, $w
    }
    open OUTPUT, ">$dat2";
    print OUTPUT "@words\n";
    close OUTPUT
}

close FILE1
```

Abbildung 3: Programmausschnitt aus *preprocess.pl*

Zuerst wird dabei durch die *chop*-Funktion jegliches letzte Zeichen des Inputs entfernt. Danach wird durch den *Transliteration Operator* *tr* und den *Flag* oder *Pattern Modifier* *d* für *delete*, all das, was zwischen den ersten beiden Schrägstrichen steht, gelöscht (siehe Christiansen et al., 2012). In diesem Fall handelt es sich um die Anführungszeichen. Analog werden Bindestriche durch Angabe eines Leerzeichens innerhalb der hinteren beiden Schrägstriche entfernt und durch dieses ersetzt. Hierauf startet eine *foreach*-Schleife innerhalb der *while*-Schleife. Sie besagt, dass jedes Element *\$w*, welches durch den *split*-Operator, der per *default* den *String* aus dem Input nach jedem Leerzeichen auftrennt, dem *Array* *@words* angehängt wird (siehe Schwartz et al., 2005). Dieser *Array* besteht somit aus einer Liste der Wörter aus *\$dat1*. Nachfolgend wird die Outputdatei geöffnet und der *Array* *@words* in diese Datei geschrieben. Zuletzt wird die Outputdatei und nach Beendigung der *foreach*- sowie der *while*-Schleife auch die Inputdatei geschlossen.

Innerhalb dieser Arbeit wurden die Outputdateien von *preprocess.pl* den zugrundeliegenden Novellen entsprechend *novelle.txt*, *venedig.txt* und *currywurst.txt* genannt. Diese Dateien können ebenfalls in Anhang 3 nachgeschlagen werden.

4.1.2 Anwendung des TreeTaggers

Im Installationspaket des *TreeTaggers*, welches, wie im vorherigen Kapitel erwähnt, im Internet auf der Seite des Instituts für Maschinelle Sprachverarbeitung der Universität Stuttgart heruntergeladen werden kann (siehe TreeTagger, 2013), ist die Installationsanleitung *INSTALL.txt* enthalten. Von dieser Anleitung ausgehend, kann man die Installation und auch die Anwendung des *TreeTaggers* im jeweiligen Betriebssystem durchführen.

Nachdem man den ersten Schritten aus der Anleitung gefolgt ist, genügt es, den Befehl *tag-german* und den entsprechenden Dateinamen anzugeben, um die Texte aus den vorbereiteten Dateien zu annotieren. Möchte man die Ausgabe nicht in der Betriebssystem-Shell stehen haben sondern in einer Datei speichern, so gibt man zusätzlich noch *>* und den gewünschten Dateinamen für den Output ein.

Zur Erstellung der annotierten Texte für die vorliegende Masterarbeit wurde für „Novelle“ zum Beispiel folgender Befehl eingegeben: *tag-german novelle.txt >nov-tag.txt*. Die POS getaggte Version von „Der Tod in Venedig“ wurde *ven-tag.txt* und jene von „Die Entdeckung der Currywurst“ *cur-tag.txt* genannt. Diese Dateien sind in Anhang 3 enthalten.

4.1.3 Perlskript zur Untersuchung des Genitivschwundes im Deutschen

Das Programm, welches zur Untersuchung des Genitivschwundes dient, ist *genitivex.pl*⁷. Das Programm zählt, wie oft bestimmte Arten von Genitiven im momentan untersuchten Text vorkommen und gibt diese danach aus. Auch die konkurrierenden Konstruktionen des Genitivs werden gezählt und dargestellt. Zum Schluss erscheint die gesamte Anzahl aller gefundenen Genitive und die relative Frequenz in Bezug auf den ganzen Text wird berechnet. Außerdem werden die Frequenzen der einzelnen Genitivtypen und auch deren Konkurrenzen bezogen auf den Gesamttext und gegebenenfalls auch auf den gesamten Genitiv errechnet.

Ähnlich wie bei *preprocess.pl* wird der User zu Beginn in der Betriebssystem-Shell über die Funktion des Programmes innerhalb der Shell informiert. Hierzu muss wiederum in das entsprechende Verzeichnis gewechselt und das Programm per *perl genitivex.pl* aufgerufen werden. Dem folgend wird der Benutzer gefragt, welche der drei Novellen untersucht werden soll. Hinzu kommt, dass dem Benutzer eine bestimmte Eingabe vorgegeben wird: Möchte er „Novelle“ untersuchen, so soll er das Stichwort *Novelle* eingeben, für „Der Tod in Venedig“ steht *Venedig* und bei Eingabe von *Currywurst* folgt die Untersuchung von „Die Entdeckung der Currywurst“. Die darauf folgende Eingabe

⁷ in Anhang 3

des Users wird per *STDIN* auf die Variable *\$name1* eingelesen und mit Hilfe von *chomp* der Zeilenumbruch am Ende der Eingabe entfernt.

Zur Kontrolle der Eingabe folgt eine *if*-Kontrollstruktur, bei der eine zu erfüllende Kondition in Klammern steht. Danach wird der Block in geschweiften Klammern ausgeführt, sobald die Kondition wahr ist (siehe Schwartz et al., 2005). Diese besagt, dass wenn *\$name1* *Currywurst* entspricht, drei weitere Variablen nach einem bestimmten Schema gefüllt werden. *\$name2* enthält die volle Namensangabe der Novelle, *\$datei1* den Reintext *timm_currywurst.txt* und *\$datei2* ist der annotierte Text *cur-tag.txt*. Die *elsif*-Kondition darunter bewirkt, dass falls *\$name1* gleich *Venedig* ist, dementsprechend für „Der Tod in Venedig“ fortgeschritten wird. Die letzte *elsif*-Kondition funktioniert auf dieselbe Art und Weise für „Novelle“. Der letzte Teil der *if*-Kontrollstruktur wird durch *else* dargestellt, welches keine Kondition in Klammern benötigt. Dieser Teil soll angewendet werden, wenn keine der obigen Eingaben getätigt wird. In diesem Fall erfolgte eine falsche Eingabe durch den User und das Programm soll nicht weiter ausgeführt werden. Dies wird mit der *die*-Funktion erreicht, welche die ihm in Klammern übergebene Nachricht zunächst ausdrückt und das Programm verlässt (siehe Schwartz et al., 2005).

Anschließend wird dem Benutzer zur Absicherung gesagt, welche Novelle er hinsichtlich des Genitivschwundes untersucht. Die gesamte Kontrollstruktur wird in dem zur Anschauung leicht modifizierten und unkommentierten Programmauszug in Abbildung 4 dargestellt. Im weiteren Verlauf sind alle Programmausschnitte ohne Kommentar aufgeführt und gegebenenfalls in ihrer Darstellung leicht verändert, ihre Funktionsweisen bleiben dabei jedoch erhalten.

```

if ($name1 eq "Currywurst") {
    $name2 = "\"Die Entdeckung der Currywurst\" von Uwe Timm";
    $datei1 = "timm_currywurst.txt";
    $datei2 = "cur-tag.txt"
}
elsif ($name1 eq "Venedig") {
    $name2 = "\"Der Tod in Venedig\" von Thomas Mann";
    $datei1 = "mann_todinvenedig.txt";
    $datei2 = "ven-tag.txt"
}
elsif ($name1 eq "Novelle") {
    $name2 = "\"Novelle\" von Johann Wolfgang von Goethe";
    $datei1 = "goethe_novelle.txt";
    $datei2 = "nov-tag.txt"
}
else {
    die("\nSie haben keine der drei Auswahlmoeglichkeiten eingegeben.
    |
    | Das Programm bricht hier ab. Bitte versuchen Sie es erneut.\n")
}

```

Abbildung 4: Dateikontrollstruktur aus *genitivex.pl*

Danach folgt die Analyse der ersten Datei. Diejenige Datei, die in der Variable *\$datei1* gespeichert ist, wird als erstes geöffnet. Kann diese nicht gefunden werden, da sie sich zum Beispiel in einem anderen Verzeichnis als das Programm befindet, so wird das Programm mit Hilfe von *die* beendet.

Zunächst sollen alle im Text vorkommenden Wörter gezählt werden, damit zum Schluss die Häufigkeit des Genitivs bezogen auf den Gesamttext berechnet werden kann. Hierzu benötigt man den ursprünglichen Reintext, der noch nicht getaggt wurde. Man könnte auch die Zeilen der *POS* annotierten Datei zählen, jedoch stehen die Satzzeichen aufgrund der Tokenisierungsart des *TreeTaggers* ebenso in einer eigenen Zeile und würden somit als Wort gezählt werden. Um dies zu umgehen, wird wie in Abbildung 5, die einen Auszug aus dem Programmcode darstellt, fortgefahren.

```
open FILE1, "$datei1"
or die("Die Datei konnte nicht geoeffnet werden oder existiert nicht.\n");

$words = 0;

while (<FILE1>) {
    chop;
    foreach $w (split) {
        $words++;
    }
}

close FILE1;
```

Abbildung 5: Ermittlung der Anzahl aller in *\$datei1* enthaltenen Wörter

Bevor eine *while*-Schleife beginnt, wird zunächst ein Zähler durch die Variable *\$words* erstellt und mit Null gleichgesetzt. Dieses Vorgehen ist in Perl im Grunde nicht nötig, wird aber zu Gunsten der Übersichtlichkeit angeführt. Die *while*-Schleife nimmt nun die erste Datei *\$datei1* als Input und entfernt per *chop* das letzte Zeichen davon. Weiterhin wird mittels einer *foreach*-Schleife für jedes Wort der Zähler durch *\$words++* um eins erhöht. Somit enthält nach Beendigung der *while*-Schleife *\$words* die Anzahl aller Wörter aus *\$datei1*. Dann wird diese wieder geschlossen.

Nun wird die zweite Datei, die den getaggt Text enthält, geöffnet. Danach wird die Variable *\$y* mit einem Zeilenumbruch gleichgesetzt, da diese das erste willkürliche Element eines Bigrammes sein soll. Wiederum wird eine *while*-Schleife, jedoch mit der zweiten Datei als Input, verwendet und das letzte Zeichen mit *chop* entfernt. An dieser Stelle taucht wieder ein alter Bekannter, der *Translition Operator*, auf. Durch ihn soll im Text *ß* durch *s* ersetzt werden.

```

open FILE2, "$datei2"
or die("Die Datei konnte nicht geoeffnet werden oder existiert nicht.\n");

$y = "\n";

while (<FILE2>) {
    chop;
    tr/B/s/;
    foreach $wd (split) {
        $z = $wd;
        if ($wd =~ /([A-Z][A-Z]+|[\$])([.][\,])/) {
            push @bigram, $x = "$y $z";
        }
        $y = $z;
    }
}

close FILE2;

```

Abbildung 6: Erstellung der aus Wörtern und *POS Tags* bestehenden Bigramme

Insgesamt soll die *while*-Schleife dazu dienen, Bigramme, die in diesem Programm aus einem Wort und einem *POS Tag* bestehen, zu erstellen und diese in einem *Array* zu speichern. Die vollständige Erstellung dieser Bigramme ist in Abbildung 6 dargestellt. Dabei wird innerhalb der *while*-Schleife durch *foreach* und *split* jedes Wort *\$wd* in einer Zeile zunächst in der Variable *\$z* festgehalten. Dann werden in einer *if*-Kontrollstruktur mit Hilfe von *Regular Expressions*, die auch reguläre Ausdrücke genannt werden, und dem *Binding Operator* *=~* jene Vorkommnisse von *\$z* gesucht, die dem Muster eines *POS Tags* entsprechen. Die regulären Ausdrücke stehen hierbei zwischen zwei nach vorne gelehnten Schrägstrichen. Der *Binding Operator* gleicht Muster auf dessen rechten Seite mit einem *String* auf der linken Seite ab (siehe Schwartz et al., 2005). Bei Schwartz et al. (2005) ist folgende Definition für reguläre Ausdrücke zu finden:

„A *regular expression*, often called a *pattern* in Perl, is a template that matches or doesn't match a given string.“

(Schwartz et al., 2005: 100)

Besteht ein Wort aus zwei oder mehreren Großbuchstaben, was durch das Muster *[A-Z][A-Z]+* erzeugt wird oder einem Satzzeichen, deren *POS Tags* durch *.* oder *,* dargestellt sind, so wird ein Paar *\$x* bestehend aus *\$y* und *\$z* dem *Array* *@bigram* hinzugefügt. Das Muster in der Form *(a/b)* drückt hierbei aus, dass entweder *a* oder *b* gesucht wird. Durch die eckigen Klammern wird die *Character Class* ausgedrückt, womit nur ein Buchstabe aus der Reichweite von *A* bis *Z* gefunden werden muss. Das *+* ist das sogenannte *Kleene Plus*, welches besagt, dass das vorausgegangene Zeichen entweder einmal oder unendlich oft vorkommt. Außerdem muss *\$* hinter einem Schrägstrich ** stehen, da *\$* sonst als regulärer Ausdruck für das Ende einer Zeile beziehungsweise das Ende eines *Strings* fehlinterpretiert werden würde (siehe

Christiansen et al., 2012). Nach Beendigung der *if*-Kontrollstruktur wird zum nächsten Wort für das folgende Bigramm übergegangen, indem $\$y$ mit $\$z$ gleichgesetzt wird. Danach schließt sich die *while*-Schleife und auch die zweite Datei wird geschlossen.

Als nächstes folgt der letzte Block des Programmcodes, der für den Benutzer in der Betriebssystem-Shell sichtbar ist. Der User wird gefragt, wie er seine Outputdatei nennen möchte. Außerdem soll die Angabe im *txt*-Format erfolgen, um anschließend mit dem Output arbeiten zu können. Wie die Datei schlussendlich genannt wird, spielt keine Rolle, es bietet sich jedoch an, die Zugehörigkeit zu den Novellen im Namen zu erwähnen, beispielsweise *nov-output.txt* für „Novelle“. Die drei Outputdateien, die innerhalb der Untersuchung in der vorliegenden Arbeit entstanden, sind ebenfalls in Anhang 3 zu finden. Wiederum wird mittels $\langle STDIN \rangle$ die Eingabe auf eine Variable, hier $\$ergebnis$, eingelesen und mit *chomp* der Zeilenumbruch entfernt. Die darauf folgende *if*-Kontrollstruktur in Verbindung mit regulären Ausdrücken und dem Binding Operator bewirkt, dass falls $\$ergebnis$ im gewünschten Format, also mit dem Kürzel *.txt* am Ende, eingegeben wird, die Bestätigung des Namens in der Shell für den User erfolgt. Gibt der Benutzer ein falsches Format an, so bricht das Programm durch Anwendung der *die*-Funktion ab. Die Ausgabe in der Betriebssystem-Shell entspricht der Darstellung in Abbildung 7.

```
C:\Users\user\Genitivschwund>perl genitivex.pl
Dieses Programm soll zur Untersuchung des Genitivschwundes im Deutschen dienen
und zaehlt den Genitiv sowie dessen Konkurrenz in einem von drei vorgegebenen
Texten.
Welchen Text moechten Sie untersuchen? Bitte geben Sie fuer "Die Entdeckung der
Currywurst" von Uwe Timm das Wort "Currywurst", fuer "Der Tod in Venedig"
von Thomas Mann das Wort "Venedig" und fuer "Novelle" von Johann Wolfgang
von Goethe das Wort "Novelle" ein <Eingabe bitte ohne Anfuhrungszeichen>.
Venedig
Sie untersuchen den Text "Der Tod in Venedig" von Thomas Mann.
Wie soll die Datei heissen, in der die Ergebnisse fuer "Der Tod in Venedig" von
Thomas Mann gespeichert werden sollen? Bitte geben Sie den Dateinamen im Format
dateiname.txt an.
ven-output.txt
Sie finden die Ergebnisse in der Datei ven-output.txt.
```

Abbildung 7: Programmaufruf und Dateiabfrage von *genitivex.pl* in der Betriebssystem-Shell

Damit die Resultate des Programmes später in die Outputdatei geschrieben werden können, muss diese zunächst unter dem zuvor angegebenen Namen mittels Variablenverknüpfung geöffnet werden. Danach werden aus den Bigrammen, die aus Wort und *POS Tag* bestehen, immer größer werdende N-Gramme erstellt. Diese stellen sozusagen die Größe der Fenster dar, in denen die Genitive auftreten. Außerdem kann man hiermit bestimmte syntaktische Folgen von Wortarten untersuchen, in denen der Genitiv auftritt und somit Muster zu dessen Erkennung definieren. So besteht beispielsweise ein Tetragramm aus zwei aufeinanderfolgenden Bigrammen, während

ein Eikosagramm, welches das größte N-Gramm darstellt, zehn hintereinander erscheinende Bigramme enthält. Die Benennung erfolgte in Anlehnung an griechische Zahlwörter.

Die N-Gramme werden innerhalb einer *foreach*-Schleife gebildet, die in Abbildung 8 zu sehen ist. Dazu wird zunächst *\$x* aus dem Array *@bigram* genommen und dem Array *@tetragram* hinzugefügt. *@tetragram* soll dabei aus zwei aufeinanderfolgenden Bigrammen bestehen, was durch *\$v = "\$w \$x"* erzeugt wird. Anschließend wird, äquivalent zur Bildung der Bigramme, *\$x* mit *\$w* gleichgesetzt, um zum folgenden Bigramm überzugehen. Die nächstgrößeren N-Gramme werden nach demselben Muster erstellt, nur dass anstatt zwei Bigramme drei oder mehrere hintereinandergeschaltet werden und dann von Durchgang zu Durchgang immer das nächste Bigramm weitergegeben wird. Danach endet die *foreach*-Schleife in Form einer schließenden geschweiften Klammer.

```
foreach $x (@bigram) {
  push @tetragram, $v = "$w $x";
  push @hexagram, $s = "$d $w $x";
  push @oktagram, $t = "$e $d $w $x";
  push @dekagram, $o = "$f $e $d $w $x";
  push @dodekagram, $p = "$g $f $e $d $w $x";
  push @tetrakaidekagram, $q = "$i $g $f $e $d $w $x";
  push @hektakaidekagram, $r = "$h $i $g $f $e $d $w $x";
  push @oktokaidekagram, $m = "$j $h $i $g $f $e $d $w $x";
  push @eikosagram, $n = "$k $j $h $i $g $f $e $d $w $x";
  $k = $j;
  $j = $h;
  $h = $i;
  $i = $g;
  $g = $f;
  $f = $e;
  $e = $d;
  $d = $w;
  $w = $x;
}
```

Abbildung 8: Erstellung größerer N-Gramme aus Bigrammen

Mit dem Ende der Herstellung dieser Bigramm-Kombinationen beginnt die Suche nach den Genitiven. Zuerst wird hierzu per *print OUTPUT* der Name der untersuchten Novelle in die Outputdatei gedruckt, damit nach Beendigung der Untersuchung noch zu erkennen ist, welche Novelle analysiert wurde. Dem folgt die Angabe, dass der erste untersuchte Genitivtyp der adnominale Genitiv ist.

Dann wird nach der ersten Gruppe der adnominalen Genitive gesucht. In diesem Fall sind das diejenigen, die einen definiten Artikel enthalten. Um deren Erscheinung später auswerten und vergleichen zu können, wird hier ein Zähler mit dem Namen *\$def* eingeführt. Nun wird als erster Schritt der adnominale Genitiv mit definitem Artikel innerhalb der Hexagramme gesucht, die aus drei Bigrammen bestehen. Hierzu wird mittels *foreach* jedes Hexagramm *\$s* aus dem Array *@hexagram* untersucht. Zu diesem Zweck wird eine *if*-Kontrollstruktur in der *foreach*-Schleife erzeugt, durch die nach bestimmten Mustern und Kombinationen gesucht werden kann.

Das erste Muster, mit dem Genitive identifiziert werden, sucht durch den *Binding Operator* und reguläre Ausdrücke nach bestimmten Hexagrammen. Diese bestehen aus einem Nomen, gekennzeichnet durch das *POS Tag NN*, gefolgt von einem definiten Artikel, also *der* oder *des*, dem *POS Tag* für einen Artikel *ART* und entweder einem Nomen *NN*, einem Eigennamen *NE*, einem substituierendem Indefinitpronomen *PIS* oder einem attribuierenden Possessivpronomen *PPOSAT*. Hierbei steht der reguläre Ausdruck *\s* für einen *Space Character* beziehungsweise ein Leerzeichen und *\b* stellt eine Wortgrenze dar. Der Punkt steht für ein beliebiges Zeichen, was in Verbindung mit dem *Kleene Star ** dazu führt, dass kein oder unendlich viele Zeichen zwischen *ART* und dem folgenden *POS Tag* im Erkennungsmuster stehen können⁸. *\$* ohne ** bezeichnet hier das Ende des *Strings* (siehe Schwartz et al., 2005).

```
if ($s =~ /NN\s\bde(r|s)\b\sART.*(NN|NE|PIS|PPOSAT)$/) {
    &defs
}
```

Abbildung 9: Suchmuster für einen adnominalen Genitiv

Abbildung 9 stellt das Suchmuster dar, so wie es im Programm zu finden ist. Wird ein derartiges Muster gefunden, so wird die Subroutine *&defs* angewendet. Subroutinen sind vom Programmierer vordefinierte Funktionen (siehe Christiansen et al., 2012), deren Definitionen das Perlskript abschließen. Diese sind besonders dann nützlich, wenn die gleichen Befehle häufiger gebraucht werden und somit nicht jedes Mal erneut eingetippt werden müssen. Die Subroutine *&defs* führt dazu, dass jedes *\$s*, das dem Suchmuster entspricht, in den Output gedruckt und der Zähler *\$def* um eins erhöht wird. Im Folgenden werden für jedes einzelne Suchmuster Subroutinen definiert und angewendet, die dem jeweiligen Genitivtyp und der momentan durchsuchten Fenstergröße entsprechen. Dabei wird immer auf die Ausgabe im Output und das Zählen der Genitive geachtet.

⁸ An dieser Stelle wäre ebenso das *Kleene Plus* ausreichend gewesen.

Anschließend wird per *elsif* das nächste Muster im Hexagramm gesucht. Dieses Muster ermittelt alle Wörter, die auf *-ere*, *-sten* oder *-ste* enden sowie als *POS Tag ADJA* tragen und in Kombination mit einem definiten Artikel sowie einem Nomen am Ende des *Strings* stehen. Das Fragezeichen *?* drückt hierbei die Optionalität des vorausgehenden Zeichens aus, weshalb *-sten?* sowohl *-ste* als auch *-sten* entspricht (Siehe Schwartz et al., 2005).

Das nächste Suchmuster gestaltet sich etwas komplexer. Zunächst wird entweder das Wort *eine*, *Eine*, *ein* oder *Ein* in Verknüpfung mit den *POS Tags ART* oder *PIS* gesucht. An dessen Stelle können aber auch das substituierende Demonstrativpronomen *die*, ein Wort mit dem Tag für Kardinalzahlen *CARD* oder die substituierenden Indefinitpronomen *keiner* oder *andere* treten. Weitere Alternativen sind *kein* mit dem Kürzel *PIAT*, *Die-* oder *denjenigen* als substituierendes Demonstrativpronomen oder schlussendlich der Artikel *dem*. Anschließend folgt ein definitiver Artikel und ein Nomen, ein Possessivpronomen, oder die Kardinalzahl *drei* am Ende des *Strings*.

Danach folgen diejenigen Suchmuster, die aufgrund von Tagfehlern oder anderen Problemen erstellt werden mussten. Beispielsweise wurde das Nomen *Haupt* fälschlicherweise vom *TreeTagger* als Eigennamen *NE* interpretiert und wird somit nicht durch das obige Muster erfasst. Außerdem müssen die Genitive, die durch *der*, eigentlich *PDS* und nicht *ART*, und *welche* eingeleitet werden, ebenfalls extra aufgeführt werden. Das *Caret ^* kennzeichnet hierbei den Beginn eines *Strings* (siehe Schwartz et al., 2005). Nach der Ausgabe dieser Genitive endet die *foreach*-Schleife.

Mit Hilfe der nächsten *foreach*-Schleife werden Genitive mit definitivem Artikel im nächstgrößeren N-Gramm, bestehend aus vier Bigrammen, gesucht. Diese Oktagramme sind im *Array @oktagram* gespeichert. Es wird in derselben Weise wie bei der Suche in den Hexagrammen vorgegangen. Auch die weitere Identifizierung des adnominalen Genitivs mit definitivem Artikel erfolgt nach demselben Schema und es wird nach Mustern, die einem solchen entsprechen, gesucht. Hierzu werden zusätzlich Dekagramme, Dodekagramme, Tetrakaidekagramme und Hektakaidekagramme benötigt.

Der nachfolgende Schritt dient zur Erfassung derjenigen Konstruktionen, die fälschlicherweise als Genitiv erkannt wurden, da sie in eines der obigen Suchmuster passen, jedoch keine genitivische Konstruktion darstellen. Hierbei wird als erstes eine Information über die Funktion der nachfolgenden Konstruktionen in die Ausgabedatei geschrieben. Anschließend wird ein Zähler *\$nogen* für diese falsch erkannten Muster eingeführt und auf Null gesetzt. Danach werden nach der gleichen Vorgehensweise wie zur Erkennung der Genitive mittels *foreach*-Schleifen bestimmte Kombinationen

innerhalb der benötigten N-Gramme gesucht. Hierbei müssen genauere Angaben über die im missinterpretierten Genitiv enthaltenen Wörter gemacht werden. Das heißt die Wörter, die in den gesuchten Konstruktionen auftreten, müssen im Suchmuster voll aufgeführt werden, damit es hier nicht zu weiteren Komplikationen kommt. Dabei muss nicht jedes einzelne Wort angegeben werden. Es kommt darauf an, ob das Wort im Einzelfall nur in einer falsch erkannten Konstruktion zu finden ist oder ob es auch in einer tatsächlich genitivischen Konstruktion aufgetaucht ist. Zwei Subroutinen *&nogens* und *&nogent* sorgen an dieser Stelle dafür, dass die Konstruktionen in die Outputdatei gedruckt und gezählt werden. Der Zähler *\$nogen* ist sehr wichtig, da damit zuerst die Angabe und Ausgabe im Output über die Anzahl der Fehl Ausgaben gemacht werden und zusätzlich die Berechnung der tatsächlichen adnominalen Genitiverscheinungen mit definitivem Artikel erfolgen kann. Dazu wird die Anzahl der falsch erkannten Konstruktionen von der Anzahl der adnominalen Genitive mit definitivem Artikel abgezogen: $\$def = \$def - \$nogen$.

Darüber hinaus wird die relative Frequenz dieser Genitivgruppe berechnet. Um diese zu ermitteln, wird die Anzahl aller adnominalen Genitive mit definitivem Artikel durch die Summe aller Wörter geteilt, was durch $\$freqdef = \$def/\$words$ erreicht wird. Ferner wird der entsprechende Prozentsatz angegeben, indem $\$freqdef$ mit 100 multipliziert und auf die Variable $\$procentfreqdef$ eingelesen wird. Nun kann der Ausdruck der Anzahl dieser Genitive sowie der berechneten Prozentsätze erfolgen. Um diese Daten im Output festzuhalten und zugleich passend zu formatieren, wird auf *printf* anstatt *print* und die *%f*-Konversion zurückgegriffen. Diese Konversion rundet die dem *String* durch Komma nachgestellten Variablen je nach Angabe auf eine bestimmte Anzahl an Nachkommastellen. So bewirkt *%8.6f*, dass die Zahl insgesamt acht Stellen besitzt, von denen sechs hinter dem Komma, beziehungsweise hier dem Punkt, stehen. Um ein Prozentzeichen in den Output drucken zu können, muss innerhalb *printf* ein doppeltes Prozentzeichen angegeben werden (siehe Schwartz et al., 2005).

Die zweite Untergruppe der adnominalen Genitive bilden in dieser Untersuchung diejenigen, die zusammen mit einem indefiniten Artikel auftreten. Erneut erfolgt zunächst die Angabe über den momentanen Untersuchungsgegenstand in der Outputdatei und es wird ein Zähler namens *\$indef* für diesen Teil des Programmes erstellt. Dabei wird das vorherige Vorgehen wiederaufgenommen. Zunächst werden dreigliedrige Wortfolgen näher analysiert, welche Konstruktionen enthalten, die einem adnominalen Genitiv mit indefinitem Artikel entsprechen. Hierbei werden die gleichen Fenstergrößen wie bei jenen Konstruktionen mit definitivem Artikel genutzt. Anschließend müssen wieder solche Konstruktionen, die fälschlicherweise als Genitiv erkannt

wurden, erfasst werden. Dabei konnte in allen drei untersuchten Texten nur ein einziges Vorkommen für diese Untergruppe gefunden werden. An dieser Stelle ist besonders wichtig, dass der Zähler dieser Konstruktionen *\$nogen* wieder mit Null gleichgesetzt wird. Ansonsten würde einfach weitergezählt werden, was zu einem falschen Ergebnis führen würde. Zur Berechnung der adnominalen Genitive mit indefinitem Artikel wird der Zähler *\$nogen* vom Zähler *\$indef* abgezogen. Somit erfolgt zuerst die Angabe über die falsch erkannten Konstruktionen und deren Anzahl, gefolgt von der Angabe über die Gesamtzahl der gefundenen adnominalen Genitive mit indefinitem Artikel einschließlich relativer Häufigkeit und Angabe des Prozentsatzes.

Die nächsten adnominalen Genitive, die näher betrachtet werden, sind diejenigen, die ein Possessivpronomen enthalten. Das bereits bekannte Vorgehen wird beibehalten und zunächst wird im Output angegeben, welche Untergruppe untersucht wird. Dann wird ein Zähler *\$poss* definiert, dem die Zahl Null zugewiesen wird. Daraufhin beginnt wieder die Suche nach Wortfolgen (inklusive der Wortart), die diesem Genitivtyp entsprechen, innerhalb des kleinsten Suchfensters, das für die Erfassung eines solchen Genitivs nötig ist. Dieses Fenster wird in dieser Untergruppe ebenfalls durch das dreigliedrige Hexagramm dargestellt. Danach wird in Okta- und Dekagrammen gesucht. Auch in den Hektakaidekagrammen, welche aus acht Bigrammen bestehen, und sogar innerhalb des *Arrays* für zehnteilige Eikosagramme können dabei Genitive mit Possessivpronomen gefunden werden. Wieder müssen falsch erkannte Konstruktionen aufgelistet und gezählt werden, um anschließend die tatsächliche Anzahl der adnominalen Genitive mit Possessivpronomen errechnen und ausgeben zu können. Dies erfolgt erneut nach demselben Schema wie bei den beiden bereits besprochenen Untergruppen.

Nachfolgend wird eine weitere Untergruppe des adnominalen Genitivs untersucht. Dabei handelt es sich um diejenigen Genitive, die in Folge eines Demonstrativpronomens auftreten. Hierzu wird in bekannter Weise ein Zähler, der als *\$dem* definiert wird, mit Null gleichgesetzt. Dann beginnt die Suche nach diesen Genitiven zunächst innerhalb der Hexagramme. Die erste Konstruktion, die gefunden werden soll, beginnt mit einem Nomen, dessen erster Buchstabe nicht *W* sein soll, da ansonsten eine falsche Konstruktion als Genitiv gewertet würde. Das Auffinden dieses Anfangsbuchstaben wird verhindert, indem innerhalb eckiger Klammern, das *Caret* \wedge vor dem Element, das nicht gematcht werden soll, eingefügt wird. Damit stellt \wedge nicht den Beginn eines *Strings* dar, sondern bewirkt eine Negation des nachfolgenden Charakters (siehe Schwartz et al., 2005). Weiterhin folgt diesem Nomen ein Demonstrativpronomen *PDAT*, welches entweder *dieser*, *dieses*, *jener* oder *jenes* lautet.

Danach soll wiederum ein Nomen, hier spielt der Anfangsbuchstabe keine Rolle, am Ende des *Strings* erscheinen. Weitere Muster, die mit einem Possessivpronomen im Genitiv zusammenhängen, werden innerhalb von Okta-, Deka- und Dodekagrammen, sowie in Tetra- und Oktokaidekagrammen identifiziert. Auch Eikosagramme müssen zur Erfassung aller Genitive mit Possessivpronomen erstellt werden. Die Aufzählung falsch erkannter Konstruktionen wird hier nicht benötigt, da an dieser Stelle keine entstanden. Weiterhin wird, analog zu den anderen Untergruppen, die gesamte Anzahl dieser Genitive in die Outputdatei geschrieben.

Nun folgt die Suche nach allen adnominalen Genitiven mit Indefinitpronomen, worüber in der Ausgabedatei informiert wird. Dazu wird der Zähler *\$inpro* gesetzt. Anschließend wird zuerst der *Array* der Tetragramme, die aus zwei hintereinander vorkommenden Bigrammen bestehen, durchsucht. Trifft das Programm dabei auf ein Muster, das ein Nomen und die substituierenden Indefinitpronomen *deinesgleichen* oder *vieler* enthält, so wird mittels *&inprov* das entsprechende Tetragramm *\$v* ausgegeben und der Zähler *\$inpro* um eins erhöht. In den Hexa- und Oktagrammen lassen sich derartige Genitive gleichermaßen finden, die mit Hilfe entsprechender Subroutinen ausgewertet werden. Bei den adnominalen Genitiven mit Indefinitpronomen gibt es keine falsch erkannten Konstruktionen, weshalb direkt die Ausgabe der Anzahl und der relativen Häufigkeit äquivalent zu den anderen Untergruppen für den User in der Outputdatei erfolgt.

Eine weitere Untergruppe der adnominalen Genitive soll mit dem Programm identifiziert werden: Die adnominalen Genitive mit Adjektiv ohne Artikel. Damit sind diejenigen Genitive gemeint, bei denen zwischen vorangehendem Nomen und dem Adjektiv kein Determinierer auftritt. Dafür wird als erstes der Zähler *\$ohneart* eingeführt. Die Suche nach diesen Genitiven erfolgt zunächst in *@hexagram*, danach wird in *@oktagram* weitergesucht. Dabei werden unter anderem solche Konstruktionen ermittelt, die einem männlichen Genitiv im Singular entsprechen. Diese Muster bestehen aus einem Nomen, gefolgt von einem attribuierenden Adjektiv mit der Endung *-en* sowie einem Nomen, dessen vorletzter Buchstabe weder *a*, *i*, *o* noch *u* sein darf und dessen letzter Buchstabe ein *s* ist. Führt man mehrere Charaktere innerhalb einer *Character Class* hinter einem *Caret* auf, so wird nicht eine Folge von Charakteren negiert, sondern die Stelle, an der diese Klasse steht, darf keinen der aufgeführten Charaktere einnehmen (siehe Schwartz et al., 2005). Diese männlichen Genitive werden mittels *&ohnearts* gezählt und im Output ausgegeben. Weitere Genitive mit Adjektiv ohne Artikel werden zudem in den *Arrays* *@dekagram* und *@dodekagram* identifiziert. Diese werden mit Hilfe entsprechender Subroutinen gezählt und in den Output geschrieben. Bei der

Erfassung dieser adnominalen Genitive mit Adjektiv entstanden einige falsch erkannte Konstruktionen, die angegeben und von der Zahl der Genitive abgezogen werden.

Als nächstes werden diejenigen adnominalen Genitive untersucht, die in vorangestellter Position erscheinen. Diese werden zum Teil durch Hexagramme erfasst. Um aber weitere vorangestellte Genitive vollständig darstellen zu können, wird ferner eine Suche innerhalb von Okta- und Dekagrammen vollzogen. Auch hierfür werden Zähler und Subroutinen aufgeführt. Abschließend werden falsch erkannte Konstruktionen ausgegeben und abgezogen, damit eine korrekte Angabe über die Anzahl der vorangestellten Genitive gemacht werden kann.

Im folgenden Abschnitt des Programmcodes wird der adnominale s-Genitiv untersucht. Dabei genügt zunächst seine Identifizierung innerhalb der Tetragramme. Zur Suche nach dem grundlegenden Muster werden Nomen, gefolgt von Eigennamen mit großem Anfangsbuchstaben und der Endung auf *-s*, mit den Tetragrammen abgeglichen. Hier müssen trotzdem innerhalb *@tetragram* noch mehrere Muster gefunden werden und die Identifizierung zweier weiterer Konstruktionen in *@hexagram* zusätzlich erfolgen. Falsch erkannte Konstruktionen konnten auch für diesen Genitivtyp ermittelt werden und müssen somit vor der Ausgabe der Anzahl der adnominalen s-Genitive noch mit-einberechnet werden.

Dann werden die aufgezählten und koordinierten Genitive gesucht, ausgegeben und gezählt. Zum Zählen wird als erstes ein Zähler, der hier *\$koord* genannt wird, benötigt. Dieser muss meistens nur um eins erhöht werden, da der erste Teil der Koordination meist schon vorher gezählt wird. Es tauchen aber auch solche Genitive auf, bei denen drei oder vier Genitive hintereinander vorkommen. Bei diesen wird der Zähler innerhalb der jeweiligen Subroutine entsprechend angepasst. Zunächst werden in den Oktagrammen alle Folgen gesucht, die aus einem Nomen, einem Eigennamen und einer Konjunktion, gefolgt von einem Nomen am Ende des *Strings*, bestehen. Diese werden mittels *&koordt* gezählt und ausgegeben. Neben der weiteren Suche von Konstruktionen im Array *@oktagram*, die einer Aufzählung oder Koordination entsprechen, werden auch längere koordinierte beziehungsweise aufgezählte adnominale Genitive in den Deka-, Dodeka-, Tetrakaideka-, Hektakaideka- und Oktokaidekagrammen gesucht. Sogar eine Suche im Array für die größten N-Gramme *@eikosagram* muss erfolgen, um alle vorkommenden Muster finden zu können. Hierbei müssen keine falsch erkannten Konstruktionen abgezogen werden und es erfolgt direkt die Ausgabe des Zählers *\$koord* in die Outputdatei.

Die letzte Gruppe, die für einen adnominalen Genitiv steht, ist jene der Relativpronomen im Genitiv. Diese ersetzen einen adnominalen Genitiv und werden deshalb als Untergruppe des adnominalen Genitivs aufgeführt. Zur Identifizierung dieser Pronomen genügt es, die Suche innerhalb der Bigramme durchzuführen, da jeweils nur nach Vorkommnissen der Pronomen *dessen* oder *deren* gesucht wird.

Im Anschluss daran wird die Gesamtzahl aller Vorkommnisse von adnominalen Genitiven berechnet. Alle Zähler, die für die Untergruppen erstellt wurden, werden zusammengezählt und die Variable *\$nom* enthält schlussendlich die Angabe über die gesamte Anzahl adnominaler Genitive des momentan analysierten Textes. Der User wird über diese Zahl in der Outputdatei informiert. Zudem wird die relative Häufigkeit des adnominalen Genitivs angegeben. Zusätzlich wird ermittelt, wie viel Prozent die jeweiligen Untergruppen bezogen auf das gesamte Auftreten des adnominalen Genitivs ausmachen.

Als nächstes werden solche Konstruktionen gesucht, die anstelle eines adnominalen Genitivs auftreten. Dabei handelt es sich zum einen um die *von*-Periphrase und die *an*-Periphrase, bei denen anstatt eines Genitivs eine präpositionale Konstruktion, eingeleitet durch *von* oder *an*, benutzt wird und zum anderen um den possessiven Dativ. Die Suchmuster für den possessiven Dativ sind hierbei ausgeklammert, da sie nur zu fehlerhaften Aussagen führen und ohnehin kein solcher Dativ in den Texten vorkommt. Entsprechend der drei Ersetzungsformen werden zunächst drei Zähler erstellt. Mit Hilfe dieser Zähler, die wiederum von Subroutinen kontrolliert werden, werden diese Genitiversetzung innerhalb der Hexagramme, sowie der Okta-, Dekä-, Dodeka-, Tetrakaideka- und sogar Eikosagramme gezählt. Zudem werden diese Konstruktionen im Output angezeigt. Anschließend erfolgt die Angabe der jeweilig gezählten Vorkommnisse für die einzelnen Formen und eine Angabe der Frequenz der Gesamtzahl dieser Ersetzungsformen des adnominalen Genitivs.

Danach wird der nächste Genitivtyp untersucht. Dies sind in diesem Fall die sogenannten freien Genitive, die sowohl prädikativ als auch adverbial vorkommen können und nicht von anderen Satzkonstituenten abhängig sind (siehe Lipavic Oštir, 2010). Wie zuvor wird eine Variable als Zähler erstellt. Danach werden all diejenigen Muster mit den Tetragrammen aus *@tetragram* abgeglichen, die für einen solchen Genitiv stehen. Dabei erscheint bei manchen Mustern das *Flag i* hinter den Schrägstrichen für den regulären Ausdruck. Dieses kleine *i* bewirkt einen *Case-insensitive Match* und sorgt dafür, dass es irrelevant ist, ob im gesuchten Muster Klein- oder Großbuchstaben erscheinen (siehe Schwartz et al., 2005). Weiterhin werden Konstruktionen für mehrgliedrige freie Genitive innerhalb der Oktokaidekagramme gesucht, bei denen das

erste Element kein Nomen sein darf. Abschließend wird die Anzahl aller erkannten freien Genitive in die Ausgabedatei gedruckt. Hinzu kommt, dass auch hier die relative Frequenz und der entsprechende Prozentsatz für diese Genitive berechnet und mitausgegeben werden.

Der Identifizierung der freien Genitive folgt die Suche nach Genitiven, die in Abhängigkeit einer Prä- oder Postposition erscheinen. Zur Untersuchung dieser wird wiederum ein Zähler erstellt und die entsprechenden Konstruktionen werden in Hexa-, Okta-, Deka-, Dodeka-, Hektakaideka-, Oktokaideka- und Eikosagrammen gesucht, um komplett dargestellt werden zu können. Direkt im Anschluss an deren Ausgabe und der Ausgabe der Gesamtanzahl sowie der relativer Häufigkeit dieser Genitive werden konkurrierende Konstruktionen ermittelt. Dabei werden Prä- beziehungsweise Postpositionen gesucht, die gewöhnlich mit Genitiv auftreten, aber in manchen Fällen mit einem anderen Kasus, hier Dativ, oder mit *von* vorkommen. Diese werden gezählt und es werden Frequenzen sowie Prozentsätze berechnet, die im Output erscheinen.

Anschließend werden Genitive gesucht, die einem bestimmten Adjektiv vorausgehen oder nachfolgen. Hierzu wird nach der Festlegung eines Zählers in den Arrays *@hexagram*, *@oktagram*, *@dekagram* und *@dodekagram* ein Adjektiv in auslösender Position für einen Genitiv identifiziert, gezählt und ausgegeben. Ähnlich wie bei den Präpositionen erfolgt danach die Ausgabe und Zählung derjenigen Adjektive, die zuvor in Verbindung mit einem Genitiv gefunden wurden, nun aber gemeinsam mit einem konkurrierenden Kasus auftreten. Zu der Identifizierung, Ausgabe und Zählung dieser Konkurrenz zum Genitiv wird nur ein einziges Suchmuster für Tetragramme benötigt. Für beide Erscheinungsformen dieser Adjektive werden wieder Gesamtanzahl, relative Häufigkeit und der daraus berechnete Prozentsatz in die Ausgabedatei geschrieben.

Als nächstes wird der adverbale Genitiv, also solche Fälle, in denen der Genitiv als Objekt eines Verbs auftritt, untersucht. Hierbei werden Teilsätze ausgegeben, damit das jeweilige Verb und der Genitiv in der Ausgabe erkennbar sind. Dies wird erreicht, indem diese Kombinationen innerhalb von Okta-, Deka-, Dodekagrammen sowie auch Tetrakaideka-, Hektakaideka- und Eikosagrammen gesucht werden. Danach wird deren Zähler, der zuvor erstellt wird, angezeigt. Parallel zu den anderen Genitivtypen wird auch für diese Gruppe ein Abschnitt erstellt, der Konstruktionen, die in Konkurrenz zum adverbale Genitiv stehen, zählt und angibt. Dazu wird nach Mustern gesucht, die ein Verb enthalten, welches als Objekt eine durch eine Präposition eingeleitete Präpositionalphrase oder eine Nominalphrase im Akkusativ hat, zuvor aber gemeinsam mit Genitiv auftrat oder normalerweise ein Genitivobjekt zu sich nimmt. Diese Muster werden mit allen N-Grammen außer den Bi- und Oktokaidekagrammen abgeglichen.

Auch hier werden abschließende Sätze, die die Anzahl dieser Vorkommnisse sowie deren Frequenz und Prozentsatz enthalten, in die Outputdatei hineingeschrieben.

Als letztes werden Konjunktionen, Adjektive und Adverbien, die einen Genitiv innerhalb des Wortes enthalten, aus den Texten extrahiert. Da es sich bei diesen jeweils nur um ein Wort handelt, wird nur der Array *@bigram* zu ihrer Identifizierung benötigt. Danach werden sie ausgegeben, sowie gezählt und deren Gesamtzahl, Frequenz und Prozentsatz werden an den Output weitergegeben.

Nun folgt die Berechnung der Anzahl aller gefundenen Genitive. Dazu werden die Zähler für alle adnominalen Genitive, freien Genitive, adverbialen Genitive, Genitive in Abhängigkeit einer Prä- beziehungsweise Postposition oder eines Adjektivs und jene von Konjunktionen, Adjektive und Adverbien, die einen Genitiv enthalten, addiert. Diese Zahl wird unter der Variable *\$total* gespeichert, die im Anschluss an diese Berechnung in die Datei geschrieben wird. Danach erfolgen zwei weitere Berechnungen: Zum einen wird die relative Frequenz des Genitivs berechnet, indem die Anzahl aller Genitive durch die Anzahl aller Wörter, die ganz zu Beginn des Programmes gezählt wurden, geteilt wird. Zum anderen wird der Prozentsatz des Genitivs ermittelt, wobei die relative Frequenz mit der Zahl 100 multipliziert wird. Anschließend wird für jeden Genitivtyp errechnet, wie viel Prozent er in Bezug auf die Gesamtanzahl der Genitive einnimmt. Dieser Prozentsatz sowie der Prozentsatz bezogen auf alle Wörter werden per *printf* im Output gespeichert.

Zum Schluss wird die Anzahl aller zum Genitiv konkurrierenden Konstruktionen im jeweiligen Text errechnet. Hierzu werden der Zähler für die *von*-Periphrase, *an*-Periphrase, den possessiven Dativ, alle Präpositionen und Adjektive mit Dativ oder *von* sowie alle Verben, die mit einem konkurrierenden Objekttyp erscheinen, zusammengezählt und auf die Variable *\$konkurrenz* festgelegt. Darüber hinaus werden die relative Häufigkeit und der entsprechende Prozentsatz der konkurrierenden Phänomene berechnet. All diese Angaben bilden den Abschluss der Outputdatei, die an dieser Stelle geschlossen wird.

Als letzter Teil des Programmes werden alle Subroutinen mittels *sub*-Kontrollstrukturen definiert, indem zunächst der gewünschte Name der Subroutinen angegeben wird und dann in die geschweiften Klammern geschrieben wird, welche Funktionen die Subroutinen annehmen sollen.

4.2 Probleme bei der Erstellung des Programmes

Eines der größten Probleme, die während der Programmierung des Perlskripts auftraten, war die Schwierigkeit, dass oftmals Wörter mit den falschen *Part-of-speech Tags* versehen wurden. Dazu kann es aus verschiedenen Gründen gekommen sein. Zum einen hatte der *TreeTagger* Probleme mit der Segmentierung bestimmter Worte, denen direkt ein Gedankenstrich folgte oder vorausging. Dieser wurde meist nicht vom Wort abgetrennt und der *Tagger* war hierdurch nicht in der Lage, die korrekte Wortart zu bestimmen. Die gleiche Problemlage tat sich bei Anführungszeichen aller Art auf. Aus diesem Grund wurde das Programm *preprocess.pl* erstellt, das die Texte von diesen Fehlerquellen bereinigt. Dabei wurden diese Zeichen komplett entfernt oder durch ein Leerzeichen ersetzt. Dies funktionierte auch bei Mehrwortlexemen, die durch Bindestriche unterteilt sind und führte nicht zu neuen Fehlerquellen.

Weiterhin gab es vor allem Tagfehler bei substantivierten Adjektiven, die eigentlich als Nomen gekennzeichnet hätten werden sollen und bei Eigennamen, die nicht als solche erkannt und als Nomen identifiziert wurden. Umgekehrt wurden einige Nomen fälschlicherweise vom *TreeTagger* für Eigennamen gehalten. Um genitivische Konstruktionen erfassen zu können, die ein fehlerhaftes Tag enthielten, mussten diese häufig einzeln, mittels einer detaillierteren Suchanfrage, aufgeführt werden, da sie oftmals nicht von einem allgemeinen Muster erfasst werden konnten. Deshalb musste zum Beispiel das „*Haupt NE des ART Eros NN*“ (ven-output.txt: Z.136), wobei *Haupt* als Eigenname gekennzeichnet wurde, bewusst extra gesucht werden. Die falsch getaggen substantivierten Adjektive konnten durch neue allgemeine Suchmuster gefunden werden.

Es kam auch vor, dass nach bestimmten Genitiven in sehr speziellen Mustern gesucht werden musste, in denen entsprechende Signalwörter enthalten sind, da ein allgemeineres Muster ebenfalls zu viele Fehler herbeiführte. Mitunter kam es nur zu einem Fehler und dieser konnte behoben werden, indem beispielsweise ein bestimmter Anfangsbuchstabe ausgeschlossen wurde oder ein Wort zwingend einen großen beziehungsweise kleinen Anfangsbuchstaben haben musste. Bei dem Muster für Eigennamen zu Beginn eines *Strings* bei der Suche nach adnominalen Genitiven mit definitivem Artikel war dies unter anderem nicht der Fall und so musste zum Beispiel „*Berlin NE der ART späten ADJA fünfziger ADJA Jahre NN*“ (cur-output.txt: Z.198) gesondert gefunden werden.

Bedauerlicherweise blieben dies nicht die einzigen Fälle, in denen falsche Wortarten zugeordnet wurden und es mussten oft einzelne Genitivkonstruktionen angezeigt werden, weil ein Artikel als Pronomen oder eine andere Wortart fehlinterpretiert wurde. Auch gab es Fehler, die bereits im ursprünglichen Text vorlagen, so ist bei Mann

„*Inbegriffleitend ADJD tätiger ADJA Jugend NN*“ (ven-output.txt: Z.664) zu finden, wobei *leitend tätiger* eigentlich von *Inbegriff* abgetrennt sein sollte. Der *TreeTagger* war deshalb nicht in der Lage ein korrektes *POS Tag* zu verteilen.

Ein anderes Problem stellte die Kodierung der Outputdatei des *TreeTaggers* dar, da die deutsche Parameterdatei in UTF-8 noch nicht von Windows unterstützt wird. Dadurch kam es dazu, dass Umlaute am Wortanfang und ß nicht korrekt identifiziert werden konnten. Diese Herausforderung wurde gelöst, indem zunächst ß während der Erstellung der Bigramme mittels *Translition Operator* durch s ersetzt wurde, um diese Wörter zu finden. Diese Lösung ist zwar etwas komplex, da dies in der Erstellung der Suchmuster berücksichtigt werden muss und die Ausgabedatei somit nicht in vollem Umfang der deutschen Rechtschreibreform entspricht, jedoch lässt sich dieses Vorgehen für diesen Zweck legitimieren. Die Ersetzung von ß darf nicht vor dem *POS Tagging* erfolgen, da ansonsten Wörter nicht richtig erkannt werden und weitere Fehler in der Wortartenzuweisung entstehen könnten.

Die Kodierung von Umlauten am Anfang eines Wortes durch den *TreeTagger* stellte insofern ein Problem dar, weil diese Wörter von den entsprechenden Suchmustern im Programm nicht gefunden werden konnten. So musste der Umlaut durch eine Kombination aus Punkt und *Kleene Star*⁹, die für kein oder unendlich viele Zeichen steht, umschrieben werden. Dabei sollte Platz für mehr als ein Zeichen sein, da die Kodierung der Umlaute voraussichtlich aus mehr als einem Element besteht. Kommt der Umlaut am Anfang des Wortes nach einem *Spacecharacter* (ls), so wurde dieser nicht einmal nach der oben genannten Umstellung erkannt. Ersetzte man jedoch ls ebenfalls durch die *Punkt-Kleene-Star*-Kombination, so wurde die entsprechende Wortfolge gefunden. Insgesamt hätte man hier wohl auch anders vorgehen können und zum Beispiel ä mit ae, ö mit oe, usw. ersetzen können. Allerdings gewährleistet die genutzte Lösung eine geringere Tagfehlerquote und führt zu einem übersichtlicheren Output. Aus unbekanntem Gründen konnte ß nicht auf die gleiche Weise ersetzt werden. Diese beiden Probleme treten möglicherweise bei Nutzung eines anderen Betriebssystems nicht auf und das Programm könnte entsprechend angepasst werden. Womöglich gelingt es auch in naher Zukunft, dass eine UTF-8 kompatible Parameterdatei für Windows erstellt wird.

Kleinere Probleme entstanden zum Beispiel vor allem dann, wenn mehrere Genitive hintereinander vorkommen, und dabei eine Fenstergröße von zehn Bigrammen überschritten wird. Da solche Fälle nicht allzu häufig auftraten, lohnte es sich nicht, noch größere N-Gramme zu erstellen. Um einen übersichtlichen Output beizubehalten,

⁹ Hier hätte auch *Kleene Plus* verwendet werden können.

wurde versucht, solche Konstruktionen, wenn möglich, auf zwei im Output aufeinanderfolgende N-Gramme zu verteilen. Das fortführende N-Gramm der beiden wurde hierbei durch das letzte Element des ersten N-Grammes am Anfang gekennzeichnet. Auch die korrekte Zählung der Genitive sollte dabei überblickt werden und die Zähler mussten dementsprechend kontrolliert und angepasst werden.

Ein eher marginales Problem ging aus der Suche nach possessiven Dativen hervor. Marginal deshalb, weil die zugrundeliegenden Suchmuster eigentlich gar nicht benötigt werden, da in keiner der drei Novellen eine Konstruktion, die tatsächlich für einen possessiven Dativ steht, gefunden werden konnte. Das Problem dabei war, dass bei der Suche Kombinationen gefunden wurden, die gar keinen possessiven Dativ darstellten, weshalb die Suchmuster auskommentiert wurden. Der Vollständigkeit halber sollten diese Suchmuster aber aufgeführt werden, denn nach Meinung einiger Autoren (siehe z.B. Fleischer & Schallert, 2011) ist die zunehmende Verwendung des possessiven Dativs ein wichtiger Indikator für den Verlust des Genitivs.

Das wohl markanteste Problem, das nicht so einfach beseitigt werden kann, entstand durch die Ambiguität weiblicher Artikel und Pronomen. So kann zum Beispiel der Artikel *der* sowohl für den weiblichen Genitiv, den männlichen Genitiv im Plural als auch für den männlichen Nominativ im Singular stehen. Dadurch kam es vor, dass Konstruktionen bestehend aus definitivem Nominativ, dem ein Nomen, welches zu einem anderen Satzglied gehört, vorausging, fälschlicherweise als Genitiv erkannt wurde. Das Suchmuster für solche Kombinationen ist jedoch grundsätzlich ein sehr robuster Indikator für adnominale Genitive und kann nicht einfach herausgenommen werden. Betrachtet man die Auftrittshäufigkeiten, wäre die individuelle Angabe der Nomen im Genitiv viel zu aufwendig. Deshalb werden die Konstruktionen, welche fälschlicherweise als Genitive erkannt werden, erneut in einer spezifischeren Eingabe, meist unter Angabe zumindest eines enthaltenen Nomens, wiederum aufgelistet und gezählt. Dabei werden sie explizit als falsch erkannt ausgewiesen. Die Zahl der falsch erkannten Konstruktionen wird anschließend wieder abgezogen und somit ist deutlich zu erkennen, wie oft ein Genitiv vorkam und welche Satzglieder zunächst falsch als solcher identifiziert wurden. Nicht nur derartige Konstruktionen mussten so ausgefiltert werden: Ferner war es notwendig andere Muster, in denen der als Genitiv erkannte Teil eigentlich zu einer nachfolgenden Satzkonstituente gehört, innerhalb der falsch erkannten Konstruktionen zu erfassen. Auch bei Eigennamen, die bereits mit einem *s* enden, musste eine solche Lösung angewendet werden. Durch diese Absätze können alle Fehler ausgebessert werden, die ansonsten nicht zu entfernen wären, ohne dass andere beziehungsweise echte Genitive darunter leiden müssen.

5. Ergebnisse der computerlinguistischen Analyse

Im vorliegenden Kapitel werden die Ergebnisse der Untersuchung, die nach Anwendung der eben beschriebenen Methoden in den Outputdateien für jede Novelle zu finden sind, besprochen. Dazu werden zuerst die Resultate dargelegt und anschließend ausgewertet.

5.1 Resultate der Genitivuntersuchung

Insgesamt zeigt die jeweilige Ausgabe des Programmes, dass bei Goethe 123 Genitive auf 7102 Wörter kommen, was einem Prozentsatz von ungefähr 1,73 % entspricht, während bei Mann 826 Genitive innerhalb von 24935 Wörtern gezählt wurden. Dies ergibt eine relative Häufigkeit von gerundet 0,033126, was circa 3,31 % entspricht. Bei Timm beträgt der Prozentsatz der Genitive hinsichtlich der Gesamtwortzahl nur 0,78 % bei 347 Genitiven in 44650 Wörtern. Dieser Prozentsatz setzt sich bei allen drei Werken aus detaillierteren Frequenzuntersuchungen zusammen, die in der Outputdatei des Programmes angezeigt werden. Hierbei wurden alle relativen Frequenzen auf sechs und die Prozentsätze auf zwei Nachkommastellen gerundet.

Bei Goethe ergibt sich bezogen auf die gesamte Wortanzahl für adnominale Genitive ein Prozentsatz von 1,46 %. Freie Genitive und Prä- beziehungsweise Postpositionen¹⁰ mit Genitiv liegen jeweils bei 0,03 %. Adjektive mit Genitiv haben gleichermaßen wie adverbale Genitive eine Auftrittswahrscheinlichkeit von 0,04 % und Konjunktionen, Adverbien und Adjektive¹¹, die Genitiv enthalten, erreichen einen Prozentsatz von 0,13 %.

Insgesamt bestehen die Genitive zu 84,55 % aus adnominalen, zu 1,63 % aus freien und zu 2,44 % aus adverbale Genitiven. Weitere 1,63 % der Genitive werden durch eine Präposition ausgelöst und 2,44 % durch ein Adjektiv. Die übrigen 7,32 % bilden die Konjunktionen, Adverbien und Adjektive.

Außerdem ergeben sich durch die Aufteilung des adnominalen Genitivs weitere Frequenzberechnungen in Bezug auf die Gesamtwortzahl. Somit machen die adnominalen Genitive mit definitivem Artikel in „Novelle“ 0,63 %, mit indefinitem Artikel 0,14 %, mit Possessivpronomen 0,07 %, mit Demonstrativpronomen 0,11 %, mit Indefinitpronomen 0,06 %, mit Adjektiv ohne Artikel 0,07 %, Voranstellungen des adnominalen Genitivs 0,07 %, der s-Genitiv 0,04 %, Koordinationen 0,23 % und Relativpronomen 0,04 % aus.

¹⁰ im weiteren Verlauf vereinfacht nur Präpositionen genannt

¹¹ im weiteren Verlauf vereinfacht nur Konjunktionen genannt

Weiterhin wurden für diese Untergruppen Frequenzprofile in Bezug auf die Gesamtanzahl der adnominalen Genitive erstellt. Dabei besteht der adnominale Genitiv in dem untersuchten Text von Goethe zu 43,27 % aus adnominalen Genitiven mit definitem Artikel, zu 9,62 % aus adnominalen Genitiven mit indefinitem Artikel, zu 4,81 % aus adnominalen Genitiven mit Possessivpronomen, zu 7,69 % aus adnominalen Genitiven mit Demonstrativpronomen und zu 3,85 % aus adnominalen Genitiven mit Indefinitpronomen. Weitere 4,81 % sind adnominale Genitive mit Adjektiv ohne Artikel, nochmals 4,81 % sind vorangestellt, 2,88 % sind s-Genitive, 15,38 % sind Koordinationen und 2,88 % Relativpronomen.

Ferner wurden diejenigen Konstruktionen, die in Konkurrenz zum Genitiv stehen, gezählt und relative Häufigkeiten für diese errechnet. Bei Goethe kommen 17 konkurrierende Konstruktionen vor, wodurch sich ein Prozentsatz von 0,24 % ergibt. Im Detail wurden zur Errechnung des Prozentsatzes Frequenzen für *von-* und *an-*Periphrasen sowie den possessiven Dativ ermittelt, wobei sich dieser Prozentsatz bei Goethe, der 0,14 % beträgt, lediglich aus der Anzahl der *von-*Periphrase zusammensetzt, da von den anderen Konstruktionen keine im Text auftraten. Hinzu kommt, dass keine Konkurrenz für Präpositionen und Adjektive mit Genitiv gefunden wurde. Hingegen wurde für Verben, die zumindest in einem der drei Texte mit Genitivobjekt auftraten, ein Prozentsatz von 0,10 % für Konstruktionen errechnet, bei denen anstatt eines Genitivs ein Präpositional- oder Akkusativobjekt gefunden wurde.

In „Der Tod in Venedig“ sind 2,97 % Prozent aller Wörter adnominale Genitive, 0,10 % sind freie Genitive oder bestehen aus Präpositionen mit Genitiv. Für Adjektive mit Genitiv beträgt der Prozentsatz nur 0,01 %. Adverbale Genitive machen in Manns Werk 0,09 % aus und Konjunktionen, die Genitiv enthalten, treten mit einer Wahrscheinlichkeit von 0,04 % auf.

Bei den adnominalen Genitiven haben jene mit definitem Artikel einen Prozentsatz von ungefähr 1,72 %, mit indefinitem Artikel 0,17 %, mit Possessivpronomen 0,30 %, mit Demonstrativpronomen 0,07 %, mit Indefinitpronomen 0,03 %, mit Adjektiv ohne Artikel 0,20 %, Voranstellungen 0,05 %, s-Genitive 0,12 %, Koordinationen 0,18 % und Relativpronomen im Genitiv erhalten einen Prozentsatz von 0,12 %.

Bezogen auf die Gesamtzahl der adnominalen Genitive machen Genitive mit definitem Artikel 57,97 % aus, jene mit indefinitem Artikel hingegen nur 5,81 %. Adnominale Genitive mit Possessivpronomen betragen 10,27 %, mit Demonstrativpronomen 2,43 %, mit Indefinitpronomen 0,95 % und mit Adjektiv ohne Artikel 6,62 %. Weiterhin erhalten

vorangestellte Genitive anteilig 1,76 %, s-Genitive 3,92 %, Koordinationen 6,22 % und Relativpronomen 4,05 %.

Der Genitiv setzt sich hier prozentual gesehen folgendermaßen zusammen: 89,59 % sind adnominal, 2,91 % frei, 2,78 % adverbial und 3,03 % werden durch eine Präposition sowie 0,36 % durch ein Adjektiv ausgelöst. 1,33 % stellen Konjunktionen, in denen Genitiv auftritt, dar.

Für konkurrierende Periphrasen konnte bei Mann ein Prozentsatz von 0,17 % ermittelt werden, der sich wiederum nur aus der *von*-Periphrase zusammensetzt. Auch Präpositionen und Adjektive mit Dativ (oder *von*) anstatt eines Genitivs, haben jeweils einen Prozentsatz von 0 %, allerdings beträgt die relative Frequenz bei den Adjektiven mit Dativ zumindest 0,000040. Konstruktionen, in denen ein Verb, welches sonst mit Genitivobjekt gefunden wurde, einen anderen Objekttyp zu sich nimmt, machen 0,03 % in Bezug auf die Gesamtwortzahl aus. Insgesamt erscheinen mit einer Wahrscheinlichkeit von 0,20 % konkurrierende Konstruktionen anstelle eines Genitivs.

In der modernen Novelle „Die Entdeckung der Currywurst“ von Timm kommt der adnominale Genitiv bezogen auf den gesamten Text insgesamt in einer Wahrscheinlichkeit von ungefähr 0,70 % vor. Dieser Prozentsatz setzt sich aus 0,42 % für adnominale Genitive mit definitivem Artikel, 0,05 % mit indefinitem Artikel, 0,06 % mit Possessivpronomen, 0,01 % mit Demonstrativpronomen und weiteren 0,01 % mit Indefinitpronomen, zusammen. Außerdem bestehen 0,03 % aus adnominalen Genitiven mit Adjektiv ohne Artikel, 0 % aus vorangestellten Genitiven, wobei die relative Häufigkeit 0,000045 beträgt, 0,05 % aus s-Genitiven, 0,02 % aus Koordinationen und 0,04 % aus Relativpronomen im Genitiv.

In der Outputdatei für diese Novelle werden wiederum die Prozentzahlen ausgegeben, welche Auskunft über die einzelnen Untergruppen des adnominalen Genitivs in Bezug auf dessen Gesamtaufreten geben. Dabei beträgt der adnominale Genitiv mit definitivem Artikel 59,55 %, mit indefinitem Artikel 7,32 %, mit Possessivpronomen 9,24 %, mit Demonstrativpronomen 1,91 % und mit Indefinitpronomen 1,27 %. Zudem sind 3,82 % der adnominalen Genitive solche mit Adjektiv ohne Artikel, 0,64 % vorangestellt, 7,64 % s-Genitive, 3,50 % Koordinationen und 5,10 % Relativpronomen.

Freie Genitive treten in Bezug auf die gesamte Wortanzahl mit einer Wahrscheinlichkeit von 0,02 % auf, während Präpositionen mit Genitiv mit einer Wahrscheinlichkeit von 0,03 % erscheinen. Adjektive mit Genitiv kamen nicht im Text vor. Adverbiale Genitive besitzen ebenfalls einen Prozentsatz von lediglich 0 % bei einer relativen Häufigkeit von 0,000022, da nur ein einziges Vorkommnis in der gesamten Novelle

gefunden wurde. Konjunktionen, Adverbien und Adjektive, die Genitiv enthalten, treten zu 0,02 % auf.

Hierbei nehmen die adnominalen Genitive einen Prozentsatz von insgesamt 90,49 %, die freien Genitive von 2,59 % und die adverbialen Genitive von 0,29 % des gesamten Genitivs im Text ein. Die durch eine Präposition ausgelösten Genitive machen 4,32 %, die durch ein Adjektiv ausgelösten Genitive 0 % und Konjunktionen, in denen Genitiv vorkommt, 2,31 % aus.

Anstelle eines Genitivs kommt im Text zu 0,23 % eine konkurrierende Konstruktion vor. Dieser Prozentsatz entsteht durch die 0,20 prozentige Wahrscheinlichkeit, dass eine *von-* oder *an-*Periphrase auftritt, sowie dem Prozentsatz von 0,01 % für Präpositionen mit Dativ oder *von*. Der Prozentsatz für Verben, die einen konkurrierenden Kasus als Objekt haben, beträgt 0,02 %.

5.2 Auswertung der erhobenen Daten

Um Aussagen über den Genitivschwund machen zu können, müssen die mit Hilfe des Programmes erhobenen Daten verglichen werden. Zuerst werden die adnominalen Genitive hinsichtlich des Genitivschwundes genauer untersucht. Hierzu werden in Tabelle 3 die unterschiedlichen Untergruppen des adnominalen Genitivs in den einzelnen Novellen einander gegenübergestellt.

Novelle	"Novelle"	"Der Tod in Venedig"	"Die Entdeckung der Currywurst"
Adnominaler Genitiv			
definitiver Artikel	0,63%	1,72%	0,42%
indefinitiver Artikel	0,14%	0,17%	0,05%
Possessivpronomen	0,07%	0,30%	0,06%
Demonstrativpronomen	0,11%	0,07%	0,01%
Indefinitpronomen	0,06%	0,03%	0,01%
Adjektiv ohne Artikel	0,07%	0,20%	0,03%
Voranstellungen	0,07%	0,05%	0,00%
s-Genitiv	0,04%	0,12%	0,05%
Koordinationen	0,23%	0,18%	0,02%
Relativpronomen	0,04%	0,12%	0,04%
gesamt	1,46%	2,97%	0,70%

Tabelle 3: Adnominale Genitive und deren prozentualer Anteil bezogen auf die gesamte Wortanzahl

In fast allen Untergruppen des adnominalen Genitivs lässt sich ein Rückgang von der ältesten Novelle bis hin zum modernen Text von Timm erkennen. Auffällig ist, dass Mann oftmals eine höhere Zahl an Genitiven als Goethe aufweist, wodurch man an dieser Stelle bereits schließen könnte, dass Thomas Mann zu einem sehr hohen Genitivgebrauch neigt. So nimmt der Prozentsatz für die adnominalen Genitive mit definitivem Artikel von Goethe zu Timm um 0,21 % ab, während er sich zu Mann beinahe verdreifacht. Die adnominalen Genitive mit indefinitem Artikel haben bei Goethe und Mann eine ähnliche Auftrittswahrscheinlichkeit, während in der modernen Novelle Timms der Prozentsatz um 0,09 % geringer ist als bei Goethe. Erstaunlicherweise nehmen die Prozentzahlen der adnominalen Genitive mit Possessivpronomen kaum ab, aber in Manns Werk ist ihr Wert deutlich erhöht. Die Genitive, die gemeinsam mit einem Demonstrativpronomen erscheinen, nehmen über den gesamten Zeitraum kontinuierlich ab. Auch bei den Indefinitpronomen ist ein ähnlicher Rückgang zu beobachten, der aber nicht ganz so markant ist. Erneut zeigt Mann in seinem Werk einen um einiges größeren Prozentsatz bei den adnominalen Genitiven mit Adjektiv ohne Artikel, während dieser aber von Goethe zu Timm abnimmt. Die vorangestellten adnominalen Genitive werden ferner im Vergleich zu Goethe bei Mann weniger und sind bei Timm so gut wie ganz verschwunden. Die s-Genitive scheinen innerhalb der

untersuchten Zeitspanne nicht zurückzugehen, denn sie zeigen wiederum nur bei Mann einen erhöhten Prozentsatz. Unerwarteterweise findet man bei Timm sogar einen leicht höheren Prozentsatz als bei Goethe. Aufgezählte und koordinierte Genitive werden bei Mann weniger und bei Timm kaum noch benutzt. Relativpronomen im Genitiv nehmen sowohl bei Goethe als auch bei Timm einen Prozentsatz von 0,04 % an, während dieser Prozentsatz bei Mann dreimal so hoch ist.¹²

Dies deutet darauf hin, dass es im Bereich der adnominalen Genitive innerhalb der gewählten Zeitspanne und der untersuchten Novellen einen deutlichen Rückgang gab. Betrachtet man auch den Prozentsatz der gesamten adnominalen Genitive in Bezug auf die Gesamtwortzahl, so geht dieser von Goethe mit 1,46 % bis zu Timm mit 0,70 % um mehr als das Doppelte zurück. Einige Genitive scheinen stabil zu sein, da sie auch im modernen Referenztext noch gleich häufig auftreten oder kaum zurückgegangen sind. Insbesondere handelt es sich dabei um die adnominalen Genitive mit Possessivpronomen, die s-Genitive und die Relativpronomen. Weiterhin ist gerade bei diesen Genitivgruppen eine auffällig erhöhte Prozentzahl bei Mann zu beobachten, was als weiteres Indiz für den verstärkten Genitivgebrauch von Thomas Mann gewertet werden kann.

Novelle	"Novelle"	"Der Tod in Venedig"	"Die Entdeckung der Currywurst"
Adnominaler Genitiv			
definiter Artikel	43,27%	57,97%	59,55%
indefiniter Artikel	9,62%	5,81%	7,32%
Possessivpronomen	4,81%	10,27%	9,24%
Demonstrativpronomen	7,69%	2,43%	1,91%
Indefinitpronomen	3,85%	0,95%	1,27%
Adjektiv ohne Artikel	4,81%	6,62%	3,82%
Voranstellungen	4,81%	1,76%	0,64%
s-Genitiv	2,88%	3,92%	7,64%
Koordinationen	15,38%	6,22%	3,50%
Relativpronomen	2,88%	4,05%	5,10%

Tabelle 4: Adnominale Genitive und deren prozentuale Anteile bezogen auf alle Genitive

Wie in Tabelle 4 verdeutlicht wird, variieren die Anteile der entsprechenden Untergruppen des adnominalen Genitivs im Laufe der Zeit und vor allem in Abhängigkeit des Autors. So wird der adnominale Genitiv mit definitivem Artikel im Vergleich zu anderen Genitiven in „Der Tod in Venedig“ und „Die Entdeckung der Currywurst“ häufiger benutzt als noch in Goethes „Novelle“. Auffällig ist hier, dass der adnominale Genitiv mit

¹² Von einer statistischen Prüfung der Signifikanz wurde insgesamt abgesehen, da die Werte der relativen Häufigkeiten für eine vernünftige Durchführung eines statistischen Testverfahrens, beispielsweise des Chi-Quadrat-Tests, zu niedrig sind.

definitem Artikel in allen drei Novellen den größten Anteil ausmacht. Dafür nimmt der Anteil des Genitivs mit indefinitem Artikel ab, bei Mann sogar stärker als bei Timm. In „Novelle“ nehmen ferner die Koordinationen einen großen Anteil ein, wenn auch einen deutlich geringeren als die Genitive mit definitem Artikel, was darauf hindeutet, dass Johann Wolfgang von Goethe gerne Genitive in aufgezählter Form verwendet hat. In „Der Tod in Venedig“ und „Die Entdeckung der Currywurst“ stehen die Genitive mit Possessivpronomen an zweiter Stelle. Außerdem nimmt der Anteil der adnominalen Genitive mit Possessivpronomen von Goethe zu Timm zu und ist auch bei Mann hoch. Diese Beobachtung verstärkt die Annahme, dass diese Untergruppe stabil geblieben ist, da sie dort Platz einnimmt, wo andere schwinden. Ein Anstieg im prozentualen Anteil von Goethe zu Timm ist auch bei den anderen beiden als stabil gewerteten Gruppen, den s-Genitiven und den Relativpronomen im Genitiv, zu erkennen. Allerdings ist hier der Prozentsatz bei Mann nicht höher als jener bei Timm.

Des Weiteren nimmt der Anteil für Genitive mit Demonstrativpronomen kontinuierlich ab. Auch die Indefinitpronomen, die anteilig gesehen bei allen drei eher selten vorkommen, nehmen ab, jedoch bei Mann stärker als bei Timm. Die Adjektive ohne Artikel bleiben in ihrem Anteil ungefähr gleich. Anders verhält es sich in Manns Werk, in dem ihr Anteil leicht erhöht ist. Auch in Bezug auf die gesamte Wortanzahl nahm der entsprechende Prozentsatz nur gering ab und Genitive mit Adjektiven ohne Artikel scheinen ebenfalls noch einigermaßen oft verwendet zu werden. Vorangestellte Genitive machen einen immer geringeren Anteil aus, was ihren Schwund verdeutlicht. Bei den aufgezählten und koordinierten Genitiven ist ein ähnlicher Rückgang zu beobachten, wobei jedoch die entsprechenden Anteile jeweils größer als jene der Voranstellungen sind. Diese deutlichen Umverteilungen lassen darauf schließen, dass viele adnominalen Genitive immer weniger verwendet werden, während einige noch stabil in ihrem Vorkommen sind.

Genitiv \ Novelle	"Novelle"	"Der Tod in Venedig"	"Die Entdeckung der Currywurst"
adnominal	1,46%	2,97%	0,70%
frei	0,03%	0,10%	0,02%
Prä-/Postposition	0,03%	0,10%	0,03%
Adjektiv	0,04%	0,01%	0,00%
adverbal	0,04%	0,09%	0,00%
Konjunktion/Adjektiv/Adverb	0,13%	0,04%	0,02%
gesamt	1,73%	3,31%	0,78%

Tabelle 5: Alle Genitivtypen und deren prozentuale Anteile bezogen auf die gesamte Wortanzahl

Tabelle 5 dient dazu, alle Genitivtypen und ihre relativen Häufigkeiten miteinander zu vergleichen. Zunächst fällt auf, dass alle Genitive zusammen bei Goethe bezogen auf die gesamte Wortanzahl 1,73 % ausmachen, bei Mann 3,31 % und bei Timm nur noch 0,70 %. Von 1828 bis 1993 ist somit der Genitivschwund deutlich erkennbar, die Prozentzahl des Auftretens des Genitivs fällt weit unter 1 %. Der Genitivanteil im gesamten Text ist bei Mann fast doppelt so hoch wie bei Goethe, obwohl „Der Tod in Venedig“ fast hundert Jahre später als „Novelle“ erschien. Deshalb wird an der Ansicht festgehalten, dass Thomas Mann in dieser Novelle einen exzessiven Gebrauch des Genitivs aufweist. Betrachtet man den höheren Gebrauch an freien Genitiven, den er im Vergleich zu beiden anderen Autoren aufweist, wird dies umso deutlicher. Der freie Genitiv, der prädikativ oder adverbial verwendet wird, muss wie der Name schon sagt, nicht zwingend benutzt werden (siehe Lipavic Oštir, 2010). Thomas Mann neigt also in „Der Tod in Venedig“ dazu, öfters Genitiv anzuwenden, als tatsächlich verlangt wird. Im Übrigen ist der freie Genitiv relativ stabil geblieben und wird bei Timm beinahe noch gleich häufig verwendet wie bei Goethe. Ein homogenes Muster lässt sich auch bei den durch Präpositionen ausgelösten Genitiven erkennen, außer dass die Prozentzahl bei Timm im Vergleich zu Goethe nicht abnimmt.

Beim adnominalen Genitiv, der den größten Anteil der Genitive ausmacht, ist eine ähnliche Verteilung wie in der Gesamtzahl zu erkennen. Wie bereits erwähnt, geht von Goethe bis Timm die Anzahl der adnominalen Genitive um circa die Hälfte zurück, während sie zu Mann um fast das Doppelte zunimmt. Bei Adjektiven, die einen Genitiv zu sich nehmen, ist allerdings schon bei Mann ein deutlicher Schwund von 0,04 % bis hin zu 0 %¹³ zu erkennen. Auch bei Timm kommt keines dieser Adjektive mehr vor. Diese Genitivart scheint im Laufe der Zeit vollends verschwunden zu sein. Der adverbale Genitiv hat im modernen Text nur noch eine Erscheinungswahrscheinlichkeit von 0,02 % und tritt dort lediglich einmal auf. Dem steht jedoch gegenüber, dass sich von Goethe diese Wahrscheinlichkeit von 0,04 % bis hin zu Mann mit 0,09 % mehr als verdoppelt. Dies deutet daraufhin, dass bereits bei Goethe der adverbale Genitiv nur sehr selten vorkam, Mann sich aber durch dessen vermehrte Verwendung auszeichnet. Im heutigen Sprachgebrauch wird der adverbale Genitiv so gut wie nicht mehr verwendet, was durch den fast vollständigen Schwund dieser Konstruktionen in „Die Entdeckung der Currywurst“ aufgezeigt werden kann. Konjunktionen, Adverbien und Adjektive, die Genitiv enthalten sind ebenfalls sehr stark zurückgegangen. Zudem ist hier ein Rückgang über Mann deutlich zu erkennen. Vermutlich sind diese Konjunktionen, Adverbien oder Adjektive zum Großteil veraltet und kommen deshalb im modernen Sprachgebrauch immer seltener vor.

¹³ bei einer relativen Frequenz von 0,000040

Anteilig gesehen werden die adnominalen Genitive bei allen drei Novellen mit Abstand am häufigsten benutzt. Tabelle 6 zeigt, dass sie dabei bereits bei Goethe einen sehr hohen Anteil von fast 85 % einnehmen, der über Mann auf bis zu ungefähr 90 % bei Timm ansteigt. Dies macht deutlich, dass schon vor fast 200 Jahren der adnominale Genitiv gegenüber dem adverbalem Genitiv dominierend war. Zudem hält diese Dominanz bis heute an und die Diskrepanz zwischen adnominalem und adverbalem Genitiv vergrößert sich sogar. Dieses Ergebnis bestätigt die These Fleischer & Schallerts (2011), dass der Genitiv im Wesentlichen ein adnominaler Kasus sei.

Genitiv \ Novelle	"Novelle"	"Der Tod in Venedig"	"Die Entdeckung der Currywurst"
adnominal	84,55%	89,59%	90,49%
frei	1,63%	2,91%	2,59%
Prä-/Postposition	1,63%	3,03%	4,32%
Adjektiv	2,44%	0,36%	0,00%
adverbale	2,44%	2,78%	0,29%
Konjunktion/Adjektiv/Adverb	7,32%	1,33%	2,31%

Tabelle 6: Prozentuale Anteile aller Genitivtypen bezogen auf alle Genitive

Interessant ist auch, dass der adverbale Genitiv bei Goethe und bei Mann einen ähnlich geringen Anteil annimmt, welcher bei Mann nur sehr geringfügig erhöht ist. Thomas Mann scheint also noch darauf bedacht gewesen zu sein, den adverbalem Genitiv zu erhalten. Bei Timm wird dieser Anteil jedoch deutlich kleiner und beträgt nur noch 0,29 %, womit der adverbale Genitiv kaum noch vorkommt.

Erstaunlicherweise nehmen freie Genitive einen höheren Anteil bei Timm als bei Goethe an. Auch bei Mann, der hinsichtlich des Gesamttexts bereits sehr viele freie Genitive benutzt hat, ist der Anteil nur gering höher als bei Timm. Dies lässt sich folgendermaßen interpretieren: Da der freie Genitiv einen höheren Anteil einnimmt, der dadurch entsteht, dass andere Genitivarten verschwinden, ist der freie Genitiv stabil und wird auch heutzutage noch häufig verwendet. Diese These wird von den zuvor dargelegten Ergebnissen gestützt. Auch Prä- und Postpositionen, die einen Genitiv auslösen, stehen anteilig in einem ähnlichen Zusammenhang, außer dass hier Manns Anteil nicht größer ist als jener von Timm. Sie scheinen somit ebenfalls stabil zu sein. Adjektive, die Genitive auslösen, machen im modernen Text keinen Anteil mehr aus und gehen bereits von Goethe zu Mann zurück. Auch Konjunktionen, Adjektive oder Adverbien, die einen Genitiv enthalten, gehen anteilig im Laufe der Zeit zurück, sie werden im Vergleich zu den anderen beiden Novellen bei Mann am wenigsten verwendet.

An dieser Stelle wird noch einmal verdeutlicht, dass einige Genitive in ihrer Verwendung über die untersuchte Zeitspanne hinweg stabil blieben, während andere offensichtlich seltener bis gar nicht mehr verwendet wurden. Adnominale Genitive wurden schon von Goethe am häufigsten benutzt und machen auch im modernen Werk den auffällig höchsten Anteil an Genitiven aus.

Weiterhin kann man beobachten, dass bei Goethe alle weiteren Genitivarten, bis auf die Konjunktionen, Adverbien und Adjektive, bei einem Anteil von ungefähr 2 % liegen. Bei Mann haben diese Typen ebenfalls einen ähnlichen Anteil, der dort circa 3 % beträgt, außer dass die Adjektive, die einen Genitiv auslösen, bei 0 % liegen. Daraus kann man schließen, dass sich Thomas Mann in der Auswahl der Genitivtypen an standarddeutschen Beispielen orientierte und außer bei den Adjektiven sowie den Genitiv enthaltenden Konjunktionen ähnliche Genitivtypen wie Goethe benutzte.

Novelle	"Novelle"	"Der Tod in Venedig"	"Die Entdeckung der Currywurst"
Genitiv & Konkurrenz			
adnominaler Genitiv	1,46%	2,97%	0,70%
Periphrasen	0,14%	0,17%	0,20%
Präpositionen mit Genitiv	0,03%	0,10%	0,03%
Präpositionen mit von/Dativ	0,00%	0,00%	0,01%
Adjektive mit Genitiv	0,04%	0,01%	0,00%
Adjektive mit Dativ	0,00%	0,00%	0,00%
adverbaler Genitiv	0,04%	0,09%	0,00%
konkurrierendes Objekt	0,10%	0,03%	0,02%
gesamte Konkurrenz	0,24%	0,20%	0,23%

Tabelle 7: Unterschiedliche Genitivarten und ihre Konkurrenz

Wie verhält sich nun der Genitiv im Vergleich zu Konstruktionen, die ihn zum Großteil verdrängt haben sollen? In Tabelle 7 sind diese konkurrierenden Konstruktionen und deren Auftrittswahrscheinlichkeit in Bezug auf den jeweils gesamten Text aufgeführt. Dabei wird voran immer die entsprechende Form des Genitivs angegeben, die ersetzt wird. In „Novelle“ treten zu 0,14 % Periphrasen auf, die einen adnominalen Genitiv ersetzen. Dies sind zusammengefasst die *von-* und die *an-*Periphrase sowie der possessive Dativ, wobei sich diese Zahl, wie bereits erwähnt, fast ausschließlich aus der *von-*Periphrase errechnet. In „Der Tod in Venedig“ erhalten diese Ersetzungsformen einen Prozentsatz von 0,17 % und in „Die Entdeckung der Currywurst“ beträgt dieser sogar 0,20 %. Im Laufe der Zeit werden also immer mehr solcher Periphrasen benutzt. Hinzu kommt, dass der adnominale Genitiv im Vergleich zu diesen Periphrasen bei Goethe noch zehnmal öfters vorkommt und bei Mann sogar 15 Mal so häufig, während bei Timm nur noch ungefähr dreieinhalb Mal häufiger ein adnominaler Genitiv als eine Periphrase auftritt. Zudem erkennt man, dass der Gebrauch von Periphrasen ansteigt, während der Gebrauch des adnominalen Genitivs abnimmt.

Somit lässt sich hier darauf schließen, dass der adnominaler Genitiv zunehmend durch Periphrasen ersetzt wurde.

Bei Präpositionen, die einen Genitiv auslösen, verhält sich das Ganze etwas anders. Während deren Verwendung über die Zeit hinweg stabil zu sein scheint, können keine beziehungsweise kaum Konstruktionen gefunden werden, bei denen Präpositionen, die grammatikalisch korrekt mit Genitiv auftreten würden (siehe Duden online, 2013), auf einen Dativ treffen. Dies ist ein weiterer Beweis für die Stabilität dieser Genitivfordernden Präpositionen. Auch Beispiele für Adjektive, die Genitiv auslösen, in den Texten aber mit Dativ oder einer Verbindung aus *von* auftraten, wurden kaum gefunden. Hier ist das kein Hinweis auf die Stabilität der genitivischen Variante, sondern betrachtet man diese, so erkennt man, dass bereits bei Mann keine solche mehr zu finden war.

Auffällig ist auch, dass gerade der Prozentsatz für diejenigen Konstruktionen, die adverbale Genitive ersetzen und in Konkurrenz zu ihnen stehen, bei Goethe höher als bei den moderneren Novellen ist. Während die adverbale Genitive, außer bei Mann, mit fortlaufender Zeit stärker zurückgehen und bei Timm nur ein einziges Vorkommen zu finden ist, nimmt auch der Anteil der konkurrierenden Konstruktionen deutlich ab. Dies kann bedeuten, dass bei bestimmten Verben zunächst noch große Konkurrenz zwischen dem Genitivobjekt und anderen Objekttypen bestand, während jedoch zunehmend die nicht genitivischen Formen benutzt werden, insgesamt diese Verben aber immer seltener bis kaum noch verwendet werden. An dieser Stelle ist zu erkennen, dass Thomas Mann sehr um den Erhalt des adverbale Genitivs bemüht ist, denn er benutzt diesen in „Der Tod in Venedig“ sogar häufiger als Goethe in „Novelle“ und auch die Anzahl der konkurrierenden Objekttypen ist in Bezug auf die Länge der Novelle geringer.

Interessant ist zudem, dass sich der Prozentsatz für alle Konkurrenzen bei allen drei Novellen im Wesentlichen nur aus den adnominalen und adverbale Konkurrenzen zusammensetzt und ungefähr gleich hoch oder vielmehr niedrig ist. Insgesamt kann damit also nicht eindeutig bewiesen werden, dass der Genitivschwund auf eine Verdrängung des Genitivs durch andere Konstruktionen zurückzuführen ist. Vielmehr scheint der Genitivschwund aus dem Rückzug des Genitivs in bestimmte Stilschichten und Register hervorzugehen (siehe Fleischer & Schallert).

Zusammengefasst zeigen die Ergebnisse, dass der Genitiv innerhalb der letzten zwei Jahrhunderte um über die Hälfte zurückgeht. Bei Thomas Mann ist zwar eine höhere relative Frequenz für den gesamten Genitiv zu finden, da er sich jedoch zeitlich

zwischen den anderen beiden Novellen befindet, wird unter anderem davon ausgegangen, dass er zu einem unverhältnismäßig häufigen Gebrauch des Genitivs neigte, zumindest was „Der Tod in Venedig“ betrifft.

Der größte Anteil der Genitive wird von den adnominalen Genitiven eingenommen, bei denen zugleich ein deutlicher Schwund zu erkennen ist. Innerhalb der adnominalen Genitive gibt es aber auch solche Typen, die über die Zeit hinweg stabil geblieben sind. Namentlich sind dies die adnominalen Genitive mit Possessivpronomen, die s-Genitive und die Relativpronomen, die für einen adnominalen Genitiv stehen. Dabei fällt auf, dass diese Genitive vorrangig Relationen ausdrücken, in denen ein Besitz angezeigt wird. Nur geringfügig schwinden die adnominalen Genitive mit definitem Artikel, die den größten Anteil der adnominalen Genitive ausmachen. Gerade diese halbwegs und ganzheitlich stabilen adnominalen Genitive benutzt Mann wesentlich häufiger als die übrigen beiden Autoren. Diese Beobachtung unterstreicht zum einen die Stabilität dieser Gruppen, zum anderen zeichnet sich dadurch wiederum Thomas Manns Hang zum Genitiv aus. Der deutlichste Schwund innerhalb dieser Gruppe ist wohl bei den vorangestellten Genitiven zu erkennen, die kontinuierlich über die untersuchten Zeitstufen verschwinden und heute fast gar nicht mehr zu finden sind. Koordinierte adnominale Genitive gehen zwar insgesamt stark zurück, hier wird aber angenommen, dass Johann Wolfgang von Goethe in seiner „Novelle“ stilistisch gesehen gerne solche Aneinanderreihungen verwendete, wobei die moderneren beiden Autoren eher vermehrt Genitive in Verbindung mit Possessivpronomen benutzten. Auffällig ist zudem, dass Periphrasen, die anstatt des adnominalen Genitivs auftreten, von den Schriftstellern im Laufe der Zeit immer häufiger verwendet werden, während der adnominale Genitiv schwindet.

Weitere stabile Genitivarten sind freie Genitive und Genitive, die durch eine Präposition ausgelöst werden. Mann zeigt für diese ebenso erhöhte Werte. Die Stabilität der von Präpositionen ausgelösten Genitive wird dadurch bestärkt, dass kaum konkurrierende Konstruktionen für sie in den drei Zeitstufen gefunden werden. Durch bestimmte Adjektive ausgelöste Genitive verschwinden schon sehr früh und sind bereits bei Mann gewissermaßen nicht mehr zu finden. Dies deutet darauf hin, dass diese Adjektive möglicherweise veraltet waren und zwischen 1828 und 1912 verloren gingen. Auch Konjunktionen, Adjektive und Adverbien mit enthaltenem Genitiv, die zwar bei Goethe noch einen relativ großen Anteil ausmachen, ereilt ein ähnliches Schicksal.

Adverbale Genitive kommen schon bei Goethe selten vor und verschwinden in den späten Werken fast ganz. Lediglich bei Mann ist wieder eine höhere Verwendung dieses Genitivtyps zu finden. Besonders interessant ist hier, dass auch konkurrierende

Objekte im Vergleich zum adverbialen Genitiv immer seltener verwendet werden. Hieraus ergibt sich die Interpretation, dass die Objekttypen zu Beginn in Konkurrenz zueinander standen, dann jedoch die nicht-genitivische Version die Oberhand gewann. Dennoch sind diese Konstruktionen mit den entsprechenden Verben heutzutage kaum noch zu finden. Diese Verben sind somit vermutlich ebenso ausgestorben oder zumindest veraltet.

Außerdem lässt sich aus dem geringen Gesamtprozentsatz aller ersetzenden Konstruktionen, der sich auch im Laufe der Zeit nicht erhöht, folgern, dass der Genitiv nicht durch die stärkere Benutzung anderer Konstruktionen verdrängt wurde. Man zeigt in den Anteilen der Genitivarten eine sehr ähnliche Verteilung zu der von Goethe. Die erhöhten Prozentzahlen bei Mann bezogen auf den ganzen Text zeichnen sich somit lediglich durch eine höhere Verwendung des Genitivs aus, auch an Stellen, an denen er nicht gefordert ist. Dies bestätigt erneut einen besonderen Hang zum Genitiv in seinem Schreibstil. Ob dies über „Der Tod in Venedig“ hinausgeht, kann durch die vorliegende Analyse jedoch nicht beantwortet werden.

Betrachtet man nun die in der Ausgabe aufgeführten Vorkommnisse der einzelnen Genitive, so erscheinen bei manchen Genitivgruppen Auffälligkeiten, welche die bisherigen Ergebnisse unterstützen. Vergleicht man die adnominalen Genitive mit definitivem Artikel in den drei Novellen, erkennt man Folgendes: Während bei „Die Entdeckung der Currywurst“ und „Novelle“ kein einziges Mal ein Genitiv mit Eigennamen auftritt, so sind in „Der Tod in Venedig“ einige zu finden. Somit zeichnet sich diese Novelle bereits hier durch ein anderes Genitivverständnis aus. Unter anderem sind diese besonderen Genitive, auch wenn sie meistens nicht vom *TreeTagger* als Eigennamen annotiert wurden, „*Haus NN des ART Aides ADJA*“ (ven-output.txt: Z.107), „*Haupt NE des ART Eros NN*“ (ven-output.txt: Z.136) oder auch „*Bleiben NN des ART trefflichen ADJA Adgios NN*“ (ven-output.txt: Z.402). Hierbei handelt es sich um griechische Gottheiten oder zumindest, nach Empfinden Aschenbachs, gottesgleiche Geschöpfe. In diesen Auszügen sticht der Bezug zur Neuklassik hervor (siehe Kurzke, 2009), der auf eine Verbundenheit zum Genitiv als Stilmittel einer gehobeneren Sprache hindeuten könnte. Bei „*Dichter NN des ART Friedrich NE*“ (ven-output.txt: Z.69) handelt es sich zwar um eine heldenhafte Sagenfigur, jedoch ist auffällig, dass keine typische Genitivendung (-s) angehängt wird. Aus dem Zusammenhang wird erkenntlich, dass hierbei das Buch und nicht die Person gemeint ist. Auch „*Insel NN der ART Adria NE*“ (ven-output.txt: Z.76) ist herauszunehmen, denn hier handelt es sich ebenfalls um eine unbelebte Erscheinung eines Eigennamens. Insgesamt sind bei allen drei Novellen eher unbelebte als belebte adnominalen Genitive mit definitivem Artikel zu finden.

Adnominale Genitive mit indefinitem Artikel zeichnen sich mit zunehmender Aktualität der Novellen lediglich dadurch aus, dass der Anteil der belebten Genitive leicht abnimmt. Somit scheinen die unbelebten indefiniten Genitive stabiler zu sein, während belebte heutzutage eher unüblich sind. Eigennamen kommen in dieser Gruppe nicht vor.

Innerhalb der adnominalen Genitive mit Possessivpronomen besteht die Auffälligkeit darin, dass hauptsächlich nur Genitive mit den Pronomen *ihrer*, *ihrer*, *seiner* oder *seiner* gebildet werden. Ansonsten erscheint lediglich *meiner* zweimal in „Die Entdeckung der Currywurst“ und einmal in „Novelle“. In „Der Tod in Venedig“ kommt neben den erstgenannten einmal das Pronomen *unserer* vor. Diese Genitive werden im modernen Text der Untersuchung zufolge immer noch in gleicher Weise verwendet wie beim ältesten, was einmal mehr die Stabilität dieser Genitivgruppe unterstreicht.

Die Demonstrativpronomen, die in adnominalen Genitiven auftreten, zeigen ein interessanteres Verhalten. Das Demonstrativpronomen *jener* bzw. *jenes* ist bei Goethe und Mann noch zu finden, während es bei Timm gar nicht mehr auftaucht. Zudem ist bei Mann ein höherer Anteil dieses Pronomens im Vergleich zu *dieser* oder *dieses* als bei Goethe zu finden. Dies verstärkt erneut die Annahme, Thomas Mann neige zu einer übermäßig häufigen Genitivverwendung, indem er bewusst veraltete Formen verwendet. Weiterhin ist eine Verbindung des Genitivs mit *aller* nur bei ihm zu finden.

Bei Goethe kommen die adnominalen Genitive mit Indefinitpronomen noch in den unterschiedlichsten Formen vor und auch die Anteile der substituierenden und attribuierenden Pronomen sind ausgeglichen. Bei den neueren Novellen fällt auf, dass jeweils nur noch ein substituierendes Pronomen auftritt. Auch bei Mann treten unterschiedliche Pronomen auf, während bei Timm fast ausschließlich *aller* gebraucht wird. Dieser Diversitätsverlust könnte an dem allgemeinen Schwund dieser Konstruktionen zum Großteil beteiligt sein.

In derselben Gruppe zeichnen sich die Genitive mit Adjektiv ohne Artikel dadurch aus, dass vor allem nur Adjektive mit der Endung *-er*, also weiblich Singular und Plural sowie männlich Plural, vorkommen. Bei Goethe und Timm ist die Parallele zu erkennen, dass jeweils nur ein einziges männliches Adjektiv im Plural auftaucht. Vergleicht man dies mit dem gesamten Vorkommen dieser Genitivsorte, so wird deutlich, dass das männliche Pluraladjektiv einen geringeren Anteil bei Timm als bei Goethe ausmacht. Bei Mann kommen diese Genitive häufiger vor, jedoch ist auch die gesamte Zahl ihres Auftretens höher. Der weibliche Genitiv scheint insbesondere hier noch gut erhalten zu sein.

Vergleicht man die vorangestellten adnominalen Genitive, so erhält man ein beachtenswertes Bild. Während in „Novelle“ nur männliche und belebte Nomina im Genitiv vorangestellt werden, treten in „Der Tod in Venedig“, bis auf eine Ausnahme, alle männlichen Voranstellungen belebt und alle weiblichen Voranstellungen unbelebt auf. Zudem werden im Vergleich insgesamt nur sehr wenige weibliche Voranstellungen gefunden. In „Die Entdeckung der Currywurst“ hingegen gibt es nur noch einen vorangestellten adnominalen Genitiv. Dieser ist weiblich sowie unbelebt. Im Rahmen der drei untersuchten Novellen ist diese Genitivart folglich nicht nur sehr markant zurückgegangen bis sogar fast verschwunden, sondern es fand womöglich eine Umkehr von männlichen belebten zu weiblichen unbelebten Genitiven statt.

Beim s-Genitiv ist auffällig, dass in allen drei Novellen im Wesentlichen nur Genitive mit Eigennamen gefunden wurden. Dies könnte der Grund für die Stabilität dieses Genitivs innerhalb des vorgegebenen Zeitrahmens sein. Die einzige Ausnahme bildet dabei „*Mengen NN Weines NN*“ (ven-output.txt: Z.778) bei Mann. Diese Konstruktion ist im Grunde genommen ungrammatisch, denn nach Plank (1995) ist der Genitiv nicht in der Lage dazu, Beziehungen des Inhalts anzugeben. Auch Meinunger (2008) würde diese Konstruktion nicht dulden, denn Massennomina können zwar ohne Artikel benutzt werden, jedoch verlieren sie dann die Fähigkeit als Genitiv realisiert zu werden.

Eine weitere Besonderheit bei Mann gegenüber den anderen Texten ist der Genitiv „*Leben NN Friedrichs NE*“ (ven-output.txt: Z.754), denn hierbei handelt es sich lediglich um den Auszug des Genitivs aus „*Leben Friedrichs von Preußen*“ (Mann, 1912: Kap.2). Zunächst erwägt man, es müsse doch *Leben Friedrich von Preußens* lauten, jedoch führt Plank (1995: 205) an: „Folgt das Attribut seinem Kopf, scheint hierbei alternativ auch der erste Namensbestandteil den Genitiv tragen zu können, besonders wenn der durch *von* angeschlossene Teil eher ein Herrschaftsbereichs- als ein Geschlechtsname ist“.

Freie Genitive fallen vor allem durch die Tatsache ins Auge, dass in der modernen Novelle nur solche vorkommen, die in Verbindung mit der Tageszeit stehen, wie „*eines ART Tages NN*“ (cur-output.txt: Z.551) oder „*eines ART Morgens NN*“ (cur-output.txt: Z.557). An dieser Stelle ist insbesondere der freie Genitiv „*Eines PIS Nachts ADV*“ (cur-output.txt: Z.555) auffällig, denn hierbei wird das weibliche Nomen *Nacht* in männlicher Genitivform verwendet. In der ältesten Novelle ist der Anteil anderer freier Genitive ausgewogen, obwohl wenige davon auftreten. In Manns Novelle hingegen ist eine sehr große Vielfalt dieser Genitivformen zu finden und auch weibliche Nomen werden in der weiblichen Genitivform angeführt. Die Form der Genitive ist hier also

nicht für die Stabilität verantwortlich, denn die Vorkommnisse in der modernen Novelle unterscheiden sich stark von jenen in den anderen beiden Texten.

Fokussiert man vergleichend die Formen der adverbale Genitive und deren Konkurrenzen in den drei Novellen, so liefert dies Aufschluss über den Rückgang dieses Genitivs. Bei Goethe kommen die Verben *bedürfen* und *bemächtigen* ausschließlich mit Genitiv vor, während *schonen* sowohl mit Genitivobjekt als auch mit Akkusativobjekt zu finden ist. An dieser Stelle nehmen auch die Verben *vergessen* und *sich erinnern* kein Genitivobjekt zu sich. Die Liste der Verben, die nur mit Genitiv vorkommen, ist bei Mann deutlich länger. Hier fällt aber auf, dass *sich erinnern* nur mit Genitivobjekt erscheint, während sein nicht-reflexives Pendant *erinnern* nur mit Präpositionalobjekt auftaucht. Auch *vergessen* ist auf beiden Seiten zu finden. In der neuesten vorliegenden Novelle findet man nur noch das Verb *bedürfen* mit Genitivobjekt. Dafür sind *erwähnen*, *erinnern* und auch *sich erinnern* mit konkurrierendem Objekttyp erschienen. Dies unterstreicht zum einen die Konkurrenzsituation, die bei einigen Verben bestand und zeigt ferner, dass viele Verben, vor allem diejenigen, die vorwiegend Genitiv fordern, im heutigen Sprachgebrauch selten auftreten. Die Verben, die heute noch häufig vorkommen, werden vorrangig mit anderen Objekttypen gebraucht.

6. Diskussion

Der allgemeine Rückgang in der Verwendung des Genitivs innerhalb der drei Novellen lässt sich nicht leugnen. Dennoch sind nicht alle Ergebnisse, die durch das Programm erzielt wurden, eindeutig als Beweise für den Genitivschwund zu werten. Geht man näher auf einige Merkmale der einzelnen Genitivtypen ein, so entstehen Zweifel, die die Gültigkeit dieser Merkmale hinsichtlich des Gesamtergebnisses betreffen.

Zwar werden bei Goethe im Vergleich zu den anderen beiden Novellen noch viele unterschiedliche Pronomen im Bereich der adnominalen Genitive mit Indefinitpronomen verwendet, doch betrachtet man diese genauer, so findet man einige Unstimmigkeiten. Beispielsweise das Pronomen *deinesgleichen*, welches als substituierendes Indefinitpronomen annotiert wurde, enthält das Possessivpronomen *deines* und ist somit weniger indefinit, sondern zeigt einen definiten Besitz an. Nichtsdestotrotz fungiert das Pronomen als ein Ganzes und ersetzt dabei eine Person und sollte nicht auf dessen einzelne Bestandteile zurückgeführt werden. Deshalb wird es innerhalb des Programmes unter den adnominalen Genitiven mit Indefinitpronomen aufgeführt. Dies ändert zwar nichts am Gesamtergebnis, der Schwund des adnominalen Genitivs mit Indefinitpronomen wäre allerdings durch die Streichung dieses nicht-eindeutigen Pronomens weniger frappierend gewesen.

Ferner ist die Vielfalt der Pronomen nicht aussagekräftig genug. Immerhin kommen bei Timm nur noch *vieler* und *aller* vor, jedoch sind die Pronomen, die bei den anderen beiden Autoren verwendet werden, durchaus noch im heutigen Sprachgebrauch zu finden. Dieser Komplettausfall einiger Indefinitpronomen in der modernen Novelle ist weniger als Grund für einen Schwund innerhalb dieser Gruppe zu deuten. Vielmehr schwindet diese Gruppe wohl im Zuge des gesamten Rückgangs des adnominalen Genitivs.

Auch das geringere Auftreten belebter Nomen bei den adnominalen Genitiven mit indefinitem Artikel könnte daraus resultieren, dass sie lediglich im analysierten Text seltener auftreten, aber nicht unbedingt allgemein zurückgegangen sind. Um dies zu evaluieren, müsste ein größerer Korpus hinzugezogen werden. Darüber hinaus ist „Novelle“ wesentlich kürzer als „Die Entdeckung der Currywurst“. Somit kann nicht hundertprozentig über das einzelne Auftreten bestimmter Formen textübergreifend argumentiert werden.

Dieser Punkt stellt auch eine Schwachstelle in der Argumentationskette für den Rückgang der adnominalen Genitive mit Demonstrativpronomen und den minimalen Schwund der Genitive mit Possessivpronomen dar. Das seltene Auftreten bestimmter

Pronomen ist eher zufällig und hat vielmehr mit der Bedeutung der Pronomen zu tun, da diese in unterschiedlichen Kontexten benötigt werden. Diese Genitivgruppen scheinen ebenfalls im Rahmen des gesamten Genitivschwundes in ihrer Zahl verringert worden zu sein, was jedoch nicht unbedingt auf die einzelnen Formen zurückgeführt werden kann. Trotzdem ist beizubehalten, dass *jener* und *jenes* heutzutage weniger gebraucht werden, womit auch die gesamte Verwendung des adnominalen Genitivs mit Demonstrativpronomen zurückgeht. Zudem wird durch deren gehäufte Verwendung in „Der Tod in Venedig“ Thomas Manns Hang zum Genitiv erneut verdeutlicht.

Eine weitere Besonderheit bei den Demonstrativpronomen weist „*Dichter NN all PIAT derer PDS*“ (ven-output.txt: Z.621) aus Thomas Manns „Der Tod in Venedig“ auf. Der vorausgehende Auszug stellt in den drei Novellen das einzige Vorkommen eines substituierenden Demonstrativpronomens dar. Allerdings ist hier nicht eindeutig, ob diese Konstruktion wirklich so besonders ist und als Indikator für Manns Präferenz des Genitivs dienen darf. Ebenso gut könnte hier nicht *derer* ausschlaggebend für die Einordnung sein, sondern *all*, welches das *POS Tag* eines attribuierenden Indefinitpronomens trägt, könnte dafür sorgen, dass die Konstruktion unter den adnominalen Genitiven mit Indefinitpronomen angeführt wird. Denn ist *all derer* nicht in etwa gleichbedeutend mit *aller*? Wiederum ist dies nur ein Einzelfall und trägt somit nicht zu einem verfälschten Ergebnis bei, dennoch sollte diesem Fall Beachtung geschenkt werden. Die Einordnung bei den Demonstrativpronomen erfolgte unter Berücksichtigung des *POS Tags* und aufgrund der Semantik von *derer*, die ebenfalls eher einem Demonstrativpronomen entspricht.

Die Argumentation über die innere Form ist bei den adnominalen Genitiven mit Adjektiv ohne Artikel hingegen deutlich nachvollziehbar. Durch Plank (1980) wird bestätigt, dass indefinite Nomen im Plural sowie weibliche Nomen ohne Artikel oder Adjektiv in einer genitivischen Konstruktion ungrammatisch werden. Insofern sind es vermutlich vor allem diese Konstruktionen, die an dieser Stelle erhalten bleiben.

Die vorangestellten adnominalen Genitive scheinen innerhalb des untersuchten Zeitraums eine Umkehr von belebten männlichen zu unbelebten weiblichen Genitiven zu vollziehen. Der allgemeine Verlust dieser Genitivart steht allerdings im Vordergrund und der Korpus reicht nicht vollständig aus, um definitiv davon ausgehen zu können, dass sich der Rückgang als ein Verlust der männlichen Variante darstellt. Bayer (2005) bestätigt, dass der Genitiv in der modernen Standardsprache fast ausschließlich in nachgestellter Form auftritt. Hierdurch wird die These unterstützt, dass der vorangestellte Genitiv generell zurückgeht und nicht aus einer Änderung der inneren Form heraus entsteht.

Ein interessanter Fall, der bereits im vorherigen Kapitel erwähnt wird, ist unter den freien Genitiven zu finden. Dort wird davon ausgegangen, dass dieser Genitivtyp stabil bleibt, sich jedoch in seiner Form stark verändert. Aber gerade diese Änderung ist erstaunlich: Während sich in den anderen beiden Novellen noch vielfältige freie Genitive finden lassen, so tauchen in der modernen Novelle nur noch Formen auf, die sich aus einem unbestimmten Artikel und einer Form der Tageszeit zusammensetzen. Wie bereits erwähnt, sticht hier „*Eines PIS Nachts ADV*“ (cur-output.txt: Z.555) heraus, da das weibliche Nomen *Nacht* als männlicher Genitiv verwendet wird. In dieser Gruppe kommen folglich nur noch feststehende Wendungen vor und es werden auch genau diese nachgeahmt, denn freie Genitive, vor allem adverbiale, sind heutzutage nicht mehr produktiv (siehe Lipavic Oštir, 2010). Aus diesem Grund ist anzunehmen, dass innerhalb dieses Genitivtyps ein deutlicher Schwund stattfand und im Gegenteil zur ursprünglichen Vermutung keinesfalls von der Stabilität des freien Genitivs ausgegangen werden kann.

Ein weiterer Problemfall wird von den Prä- und Postpositionen, die Genitiv übertragen, verdeutlicht. Oftmals ist nicht klar, ob es sich beim Attribut der Prä- oder Postposition tatsächlich um einen Genitiv handelt. Bereits die beiden Vorkommnisse, die bei Goethe gefunden wurden, lassen hierbei Zweifel aufkommen. Aufgrund der Ambiguität des Artikels *der* kann nicht eindeutig unterschieden werden, ob ein Genitiv oder ein Dativ auftaucht. Konsultiert man dazu die Experten vom Duden¹⁴, wird schnell deutlich, dass die Präposition *anstatt* vorrangig mit Genitiv auftritt und an dieser Stelle richtig erkannt wird. Auch die Konstruktion mit *wegen* kann als Genitiv interpretiert werden, eine dativische Interpretation wäre als umgangssprachlich zu werten. Derartige Konstruktionen lassen sich zudem bei Mann finden, können aber nach Konsultation des Dudens ebenfalls als Genitive gezählt werden. Eine interessante Konstruktion bei Mann, die wiederum bei Timm dreimal auftaucht, ist jene, die aus *voller* und einem nachfolgenden Nomen ohne Artikel besteht. Wie kann man sich hier sicher sein, dass es sich um einen Genitiv handelt? Der Duden gibt an, dass die Präposition *voller* nur in seltenen Fällen gemeinsam mit Dativ auftritt, weshalb die genitivische Lesart bevorzugt werden sollte. Weiterhin fallen bei Timm mehrere Konstruktionen mit *wegen* auf, bei denen unklar ist, ob Dativ oder Genitiv gebraucht wurde.

Geht man davon aus, dass es sich bei Novellen um Textformen handelt, die wegen ihres gesprächshaften Charakters nahe an der gesprochenen Sprache liegen, so könnte man durchaus behaupten, dass die gefundenen Problemfälle von dativischer Natur sind. Zudem könnte man postulieren, dass eine Umkehr von Genitiv zu Dativ

¹⁴ Erscheint der Begriff Duden, wird stets auf Duden online (2013) zurückgegriffen.

stattfand und während in den älteren Novellen solche Konstruktionen noch mit Genitiv vorkamen, sie in der modernen Novelle bereits mit Dativ auftreten. Unterstützt wird diese These von den Ergebnissen für die Präpositionen, die eigentlich Genitiv zu sich nehmen sollten, aber mit einem anderen Kasus erscheinen. Diese sind schließlich bei den älteren beiden Novellen nicht zu finden, während in Timms Werk drei derartige Konstruktionen auftreten. Laut Duden handelt es sich bei „wegen APPR Fahnenflucht NN“ (cur-output.txt: Z.591) um ein allein stehendes, stark gebeugtes Substantiv im Singular, welches in Verbindung mit *wegen* nicht gebeugt wird. Zwar kann an dieser Stelle weder von einem Dativ, noch von einem präpositionalen Objekt gesprochen werden, jedoch taucht auch kein Genitiv auf. Des Weiteren kommt *wegen* sogar einmal mit Dativ *vor* und *inmitten* mit einem Präpositionalobjekt. Somit kann die Stabilität der Genitive, die von einer Präposition ausgelöst werden, stark angezweifelt werden.

Nichtsdestotrotz werden auch bei Timm eindeutige Beispiele für die genitivische Verwendung derselben Präpositionen gefunden, zum Beispiel „wegen APPR des ART Currys NN“ (cur-output.txt: Z.573), und die Konkurrenzen müssen von anderen Faktoren abhängig sein. Beispielsweise könnte der dialektale Bezug des Sprechers oder Erzählers dafür verantwortlich sein. Auffällig ist, dass sobald eine nicht-eindeutige oder konkurrierende Konstruktion gefunden wird, Frau Brücker das Wort im Gespräch übernimmt. Dies unterstreicht den umgangssprachlichen Charakter dieser Konkurrenzen, denn Frau Brücker neigt zu einem Hamburger Dialekt (siehe Timm, 2008). Um also genaue Aussagen über den Kasus mancher Konstruktionen machen zu können, müsste hier eine neue Untersuchung unter Berücksichtigung der jeweiligen Sprecher und Gesprächslage angesetzt werden. Die Behauptung, dass dieser Bereich des Genitivs im Laufe der untersuchten Zeitspanne komplett stabil geblieben sei, muss zurückgenommen werden.

Über Adjektive, die Genitive auslösen, kann keine äquivalente Aussage gemacht werden, jedoch lassen sich auch hier Schwachstellen aufzeigen. Für die drei Adjektive, die bei Goethe gefunden wurden, konnte nicht bewiesen werden, dass es sich tatsächlich um Genitive handelt, denn wiederum trat der ambige Artikel *der* auf. In „Der Tod in Venedig“ kommt dreimal ein Genitiv vor, der durch das Adjektiv *bewußt* ausgelöst wurde. Hentschel (2010b) führt dieses Adjektiv eindeutig als Genitiv auslösend auf. In der neuesten Novelle wird kein derartiges Vorkommen registriert. Insgesamt wurde nur eine einzige konkurrierende Konstruktion mit eindeutigem Dativ bei Mann gefunden, wofür jedoch dort kein genitivisches Gegenstück vorkommt. Gänzlich kann also nicht von einem Genitivschwund dieser Formen ausgegangen werden, denn

die Eindeutigkeit der Funde bei Goethe kann nicht bewiesen werden und auch die konkurrierenden Fälle geben keinen Aufschluss hierüber.

Bezieht man sich, bei näherer Betrachtung der adverbale Genitive, auf die in der Auswertung beschriebene Konkurrenzsituation, so sollte man beachten, dass über den Partivitätseffekt und den Verlust der Aspektkategorie hier nicht mehr zu argumentieren ist. Denn wie im Abschnitt über die Ursachen für den Genitivschwund im zweiten Kapitel bereits aufgeführt wurde, fanden diese Phänomene bereits in der althochdeutschen beziehungsweise mittelhochdeutschen Zeit statt. Allerdings könnte die sekundäre Begleiterscheinung dessen – der Rückzug einiger adverbale Genitive in bestimmte gehobene Stilschichten – für die unterschiedliche Benutzung der Verben verantwortlich sein. Die adverbale Genitive bei Goethe sind alle eindeutig und auch der einzige adverbale Genitiv, der bei Timm gefunden wurde, kann als solcher gewertet werden, denn Hentschel (2010a) führt den Genitiv als Objekttyp des Verbs *bedürfen* auf. Auch die vielen adverbale Genitive, die in „Der Tod in Venedig“ identifiziert wurden, sind zum Großteil eindeutig und bei den nicht-eindeutigen Fällen gibt es wenig Zweifel daran, dass diese eigentlich ein konkurrierendes Objekt darstellen. Auffällig ist, dass *sich erinnern* in Verbindung mit Genitiv steht, während *erinnern* mit Präpositionalobjekt erscheint. Bei Goethe hingegen kam *sich erinnern* bereits in konkurrierenden Konstruktionen vor. Betrachtet man *sich erinnern* bei Mann näher, so wird deutlich, dass im Gegensatz zu *erinnern*, *sich erinnern* mit Genitiv dann verwendet wird, wenn von den persönlichen positiven Erinnerungen Aschenbachs die Rede ist. Wird von negativen Erinnerungen gesprochen, so kommt *erinnern* mit Dativ zum Einsatz (siehe Mann, 1912). Die Verben stehen also in Konkurrenz und ferner in einer anderen Verteilung als zuvor, um als Stilmittel zu dienen und positive Empfindungen auf eine höhere Stilebene zu heben. In der modernsten untersuchten Novelle kommt *erinnern* sowohl reflexiv als auch nicht-reflexiv nur in konkurrierenden Formen zum Genitiv vor. Beim Verb *schonen*, welches in „Novelle“ sowohl mit Genitiv als auch Akkusativ auftritt, ist der Sprecherbezug nicht ganz deutlich. Man kann aber feststellen, dass die Wärterin und der Honorio lediglich Akkusativ benutzen, während das einzige Genitivvorkommen in Abhängigkeit von *schonen* nur beim Fürsten in einem Konditionalsatz zu finden ist (siehe Goethe, 1828). Dies deutet auch auf den Genitiv als Stilmittel für die gehobene Sprache des Fürsten hin, was allerdings durch die zusätzliche Benutzung desselben Verbs mit Akkusativ vom Fürsten abgeschwächt wird.

Hinsichtlich der Frequenzen des adverbale Genitivs wird deutlich, dass dieser heutzutage kaum noch verwendet wird, weil er von Goethe zu Timm bis auf 0 % sinkt. Da

auch hier Mann einen erhöhten Prozentsatz aufzeigt und deutlich andere Verben als Goethe benutzt, kann für die übermäßige Benutzung des Genitivs bei Thomas Mann argumentiert werden. Über das Aussterben bestimmter Verben können an dieser Stelle keine sicheren Aussagen gemacht werden, es besteht jedoch die Tendenz dazu, dass die Diversität der Verben, die ein Genitivobjekt zu sich nehmen, geringer wird. Dies wird primär beim Übergang von Mann zu Timm deutlich.

Unerwarteterweise ist die Prozentzahl für die konkurrierenden Objekte bei Goethe am höchsten und nimmt in den späteren untersuchten Werken nur sehr gering ab. Fokussiert man aber die absoluten Zahlen, so ist die Zahl bei Timm am höchsten, allerdings macht dieser Wert bezogen auf die Gesamtwortzahl am wenigsten aus. Erwartungsgemäß sollte zu sehen sein, dass immer mehr ersetzende Konstruktionen im Laufe der Zeit benutzt werden. Dieser vermeintliche Fehler kann auf die Datenerhebung zurückgeführt werden. In der vorliegenden Arbeit wurden all diejenigen Verben, die in allen Texten auch mit Genitiv gefunden wurden, dahingehend überprüft, ob sie ebenso mit einem konkurrierenden Objekt auftraten und dann gezählt. Diese Zählweise ist in gewisser Weise ungünstig, jedoch muss an einer bestimmten Stelle ein Rahmen zur Erfassung dieser Verben definiert werden, denn alle Verben mit anderen Objekten zu zählen, wäre noch viel irreführender. Es konnten auch nicht nur diejenigen Verben mit Genitiven, die spezifisch in der jeweiligen Novelle auftraten, nur in eben dieser Novelle auf einen konkurrierenden Kasus übertragen werden. Dann hätten Details, wie im Falle von *erinnern* versus *sich erinnern* bei Mann, nicht herausgearbeitet werden können. Es ist also schwierig, hier den richtigen Maßstab zu finden. Vergleicht man die absolute Anzahl der adverbale Konkurrenz mit der absoluten Zahl der adverbale Genitive, so wird deutlich, dass bei Goethe die adverbale Genitive in einem Verhältnis von 3:7 in Bezug auf die Konkurrenz vorkommen. Es herrschte also bereits im frühesten Werk ein eher niedriger Gebrauch des adverbale Genitivs. Bei Mann beträgt dieses Verhältnis 23:7, was mit der exzessiven Verwendung des Genitivs erklärbar ist. Timm zeigt in dieser Kategorie ein Verhältnis von 1:10. Geht man näher auf diese Verhältnisse ein, so kann für einen Rückgang des adverbale Genitivs bei gleichzeitigem Anstieg der Konkurrenz argumentiert werden. Diese Schwierigkeiten in der Erfassung ergaben sich auch bei der Zählung der Konkurrenzen von genitiv-auslösenden Adjektiven und Präpositionen.

Zusätzlich sollte erwähnt werden, dass die konkurrierenden Konstruktionen keine eindeutige Auskunft über den Genitivschwund geben können. Wider Erwarten ist die relative Häufigkeit aller Konkurrenzen bei den drei Novellen gleich hoch. Dies kann zum einen daran liegen, dass nur diejenigen Konstruktionen bei adverbale Genitiven

sowie genitivauslösenden Adjektiven und Prä- oder Postpositionen als Konkurrenz erfasst wurden, die zuvor bereits mit Genitiv gefunden wurden. Weiterhin ist auch kein sogenannter possessiver Dativ ermittelt worden, obwohl die Konstruktion nicht erst im modernen Deutsch auftritt, sondern bereits seit längerer Zeit dabei ist, den Genitiv zu verdrängen (siehe Meinunger, 2008). Es ist denkbar, dass der possessive Dativ eher im süddeutschen Sprachgebrauch vorkommt und deshalb in der modernen Novelle nicht auftaucht, da Frau Brücker aus Hamburg stammt (siehe Timm, 2008). Eine eindeutige Erklärung ist dies aber nicht, denn Fleischer & Schallert (2011) geben ein Beispiel für den possessiven Dativ im Nordniederdeutschen an.

Es konnte zwar der Verfall des Genitivs mit Hilfe der Untersuchung gezeigt werden, allerdings steht dieser scheinbar nicht in Zusammenhang mit einer Verdrängung durch konkurrierende Konstruktionen. Die relative Häufigkeit der Periphrasen, die den adnominalen Genitiv ersetzen, nimmt im zeitlichen Verlauf der Novellen zu, es ist aber nicht einfach festzustellen, ob diese Periphrasen nur eine ersetzende Funktion haben oder an gewissen Stellen sogar die einzig richtige Variante sind, um eine bestimmte Relation auszudrücken. Denn findet man „kein beugsames Element, weder Artikel noch Adjektiv, in der Nähe, wie das bei indefiniten Konstruktionen bisweilen der Fall ist, dann muss man das Element von einschieben“ (Meinunger, 2008: 75). Somit sind zwei entgegengesetzte Schlussfolgerungen möglich: Zum einen könnte der adnominalen Genitiv dadurch dezimiert werden, dass er durch *von*- oder seltener *an*-Periphrasen ersetzt wird, wobei die anderen konkurrierenden Konstruktionen über die adnominalen Genitive hinaus nur einen geringen Anteil im Rahmen der Verdrängung des Genitivs ausmachen. Zum anderen kann argumentiert werden, dass diese Konstruktionen den Genitiv eigentlich nicht ersetzen, sondern an gewissen Stellen benutzt werden müssen. Diese grammatisch benötigten *von*-Periphrasen können somit den Genitiv nicht verdrängen, denn sie werden schon immer gebraucht. Betrachtet man die Periphrasen, so wird deutlich, dass vor allem bei Timm in den meisten Fällen ein beugsames Element vorhanden ist. Ist dies nicht der Fall, so beinhaltet die *von*-Periphrase fast nur Eigen- und Städtenamen und ersetzt hier den s-Genitiv. In „Der Tod in Venedig“ ist hingegen zu sehen, dass die meisten *von*-Periphrasen, solche sind, die an dieser Stelle stehen müssen. Auch bei Goethe sind diese zu finden, jedoch ist die Zahl der ersetzenden *von*-Periphrasen geringfügig größer. Somit lässt sich ein Anstieg in den Ersetzungen feststellen. Außerdem ergibt sich, dass Mann weniger tatsächliche Ersetzungen des Genitivs benutzt als die beiden anderen Autoren. An dieser Stelle müsste eine detailliertere Untersuchung der Periphrasen angesetzt werden und genauer überprüft werden, inwiefern diese den Genitiv beeinflussen und welche Funktion sie tragen, damit eindeutige Ergebnisse vorgelegt werden können.

Eine weitere Schwachstelle der Datenanalyse liegt in der Erfassung der Konjunktionen, Adjektive und Adverbien, die Genitiv enthalten. Diese stellen nicht eindeutig Genitive dar, sie sind entweder veraltete Formen adnominaler Genitive (siehe Duden online, 2013) oder beinhalten einen Genitiv in beispielsweise pronominaler Form. Man könnte sie stellenweise den freien Genitiven zuordnen oder auch den Pronomen, jedoch unterscheiden sie sich von den restlichen Vertretern dieser Gruppen. Deshalb ist es sehr schwierig, sie einer anderen Kategorie unterzuordnen. Auch ob sie überhaupt erfasst werden sollen, ist zweifelhaft, sie machen zum Beispiel bei Goethe einen erheblichen Anteil der Genitivverwendungen aus und könnten die Ergebnisse verfälschen. Nichtsdestotrotz werden sie erfasst, weil sie auf irgendeine Art und Weise eine genitivische Relation ausdrücken oder zumindest früher einmal die Funktion eines Genitivs besaßen. Zudem wird deutlich, dass diese Vorkommen veraltet sein müssen, denn sie sinken rapide ab. Hundertprozentig lässt sich dies jedoch nicht entscheiden, denn die Aufttrittsformen dieser Kategorie sind in den einzelnen Novellen sehr unterschiedlich und man könnte argumentieren, dass das zeitliche Absinken eher mit der nur vagen Definiertheit der Kategorie zu tun hat oder die Verwendung dieser stilistische Gründe hat. Der Vollständigkeit halber sollten diese Vorkommnisse trotzdem aufgeführt werden. Möchte man sie dennoch nicht mitzählen, so kann man im Programmcode den Zähler dafür entfernen und sie spielen für das Gesamtergebnis keine Rolle mehr.

Eine zusätzliche Schwierigkeit stellt die Art und Weise dar, in der aufgezählte und koordinierte Genitive aufgenommen werden: Wie oft sollte man diese zählen? Sind zwei koordinierte Genitive eine Einheit und deshalb als ein Genitiv zu zählen oder sind es doch zwei eigene Genitive? In dieser Untersuchung wurden Koordinationen und Aufzählungen folgendermaßen aufgefasst: Die Konjunkte, zumindest das zweite, können nicht alleine stehen oder aus der Konstruktion entfernt werden, denn dann können sie die gewünschte Relation nicht mehr ausdrücken. Somit bilden sie zwar eine untrennbare Einheit, trotzdem beziehen sich zwei Genitive auf denselben Referenten, denn zum Beispiel *das Recht des Vaters und der Mutter* ist sowohl das Recht der Mutter als auch das Recht des Vaters. Ungünstigerweise gibt es hier aber eine weitere Lesart, denn das Recht kann auch bei dem Vater und der Mutter gleichermaßen liegen, diese bilden eine Einheit und das Recht bezieht sich auf beide. Syntaktisch vereinfacht unterscheiden sich die Lesarten wie folgt:

zwei Genitive: [*das Recht [des Vaters]_{NP}[und der Mutter]_{NP}]_{NP}*
ein Genitiv: [*das Recht [des Vaters und der Mutter]_{NP}]_{NP}*

Zur einheitlichen Erfassung all dieser und ähnlicher Konstruktionen musste eine Entscheidung getroffen werden. Es wurde festgelegt, jeden Genitiv einzeln zu zählen

und die Koordinationen sowie Aufzählungen ebenfalls aufzuführen, da diese Genitive erneut gewichtet werden sollten. Äquivalent dazu werden auch Periphrasen, die aus zwei koordinierten Elementen bestehen, gezählt.

Bei den Koordinationen und Aufzählungen ist wiederum zu erkennen, dass bei Goethe ein deutlich höherer Prozentsatz für diese Konstruktionen errechnet wird, der im Laufe des untersuchten Zeitraums stetig abnimmt. Dies kann entweder aus dem allgemeinen Schwund des adnominalen Genitivs resultieren oder Goethe zeigt eine besondere Vorliebe für diese Konstruktionen. An dieser Stelle werden der hohe Anteil, der sich bei Goethe im Vergleich zum Gebrauch der anderen adnominalen Genitivarten zeigt und die Tatsache, dass auch Manns Prozentsatz und Anteil geringer sind, so interpretiert, dass sich dieses gehäufte Vorkommen bei Goethe auf seinen Schreibstil zurückführen lässt und zusätzlich eventuell in geringem Maße von dem allgemeinen Genitivrückgang beeinflusst wird.

Insgesamt ist es formal gesehen schwierig, die Novellen einzuordnen und ihre Nähe zur gesprochenen Sprache zu klassifizieren. Das Gespräch, das in der Novelle im Mittelpunkt steht, sollte eine hohe Affinität zur Umgangssprache haben. Dafür konnten jedoch nur selten Hinweise gefunden werden, wie im Falle von Frau Brücker in „Die Entdeckung der Currywurst“. Sonstige Genitivverluste oder Ersetzungen lassen sich nur schwer auf solche Texteigenschaften und gattungstechnische Kriterien zurückführen. Nichtsdestotrotz vertreten die Novellen die gegenwärtige Sprache ihrer Zeit und erlauben somit Aussagen über den Genitivgebrauch.

Weiterhin problematisch sind die unterschiedlichen Längen der Novellen, denn „Novelle“ ist wesentlich kürzer als die anderen beiden und sogar „Die Entdeckung der Currywurst“ ist, gemessen an der Wortzahl, fast doppelt so lang wie „Der Tod in Venedig“. Hier könnten bessere und präzisere Ergebnisse erzielt werden, würde man annähernd gleich lange Novellen aussuchen. Zusätzlich stellt sich die Frage, ob Timms Werk dem Novellenbegriff noch gerecht wird und sein Buch nicht die kürzere bis mittlere Länge überschreitet. Zudem wurde eingangs festgestellt, dass „Der Tod in Venedig“ keine mustergültige Novelle darstellt (siehe Aust, 2012). Dem ist jedoch entgegenzusetzen, dass nicht absolute, sondern relative Zahlen in der vorliegenden Untersuchung zur Evaluierung benutzt wurden und würde die Länge tatsächlich relevant sein, so hätte sich kein deutlicher Genitivschwund von alt zu neu abgezeichnet. Die Aussagen über die genauen Formen der Genitive wirken durch die ungleichen Korpora weniger fundiert. Die Länge von Timms Novelle ist den Charakter betreffend nicht ausschlaggebend, denn andere typische Novellenmerkmale, wie die

unerhörte Begebenheit, die Rahmenerzählung und das Gespräch sind eindeutig vorhanden.

Der abweichende Novellencharakter bei Mann könnte verantwortlich dafür sein, dass der Autor fast in jeder Kategorie deutlich mehr Genitive benutzt als die anderen, da er weniger nah an der Alltagssprache und näher an der gehobenen Sprache schreibt. Dies sei jedoch dahingestellt, denn die Novellenauszeichnung seines Werks stammt von ihm selbst und sein hoher Genitivgebrauch scheint wohl ebenfalls von ihm gewollt. Zudem kommen bei ihm durchaus umgangssprachliche Konstruktionen vor, denn es wird häufig die indefinite Lesart von *welche* gefunden, die als umgangssprachlich gilt (siehe Meinunger, 2008).

Der Bezug zur Neuklassik, der in allen drei Novellen vorzufinden ist, ist bei Thomas Mann am stärksten. Nicht nur die häufigen Bezugsnahmen zu griechischen Göttern und Helden sowie Philosophen sind auffällig, sondern auch sein Schreibstil ist in dieser Hinsicht als neuklassizistisch zu deuten. Vor allem im 18. und 19. Jahrhundert waren Griechisch und Latein sehr angesehen und griechische Schreibweisen wurden adaptiert (siehe Meinunger, 2008). So erscheint bei Mann das Wort *Stil* in „*Meisterhaltung NN unseres PPOSAT Styls NN*“ (ven-output.txt: Z.592) in einer dem Griechischen nahegelegten Schreibweise. Die hier besprochene Novelle von Thomas Mann erschien zwar erst im 20. Jahrhundert, dennoch wird deutlich, dass er sich gerne an veralteten Normen orientiert. Dieser Rückbezug zur Antike könnte auch maßgeblich für seinen hohen Genitivgebrauch sein. Bei Goethe ist ein derartiger Bezug zur Antike oder zum Klassizismus nur in den lyrischen Abschnitten seiner Novelle zu erkennen, jedoch entstand „Novelle“ im Zeitrahmen der Klassik (siehe Aust, 2012). Bei Timm kann eine neuklassische Deutung nur auf interpretatorischer Ebene stattfinden.

Thomas Mann wird häufig an deutschen Schulen unterrichtet mit dem Zweck, den Schülern seinen Stil näher zu bringen (siehe Kurzke, 2009). Es ist allerdings Goethe, der größte deutsche Dichter, der als Instanz der Normierer gewertet wird, womit vor allem die Vertreter des Dudens gemeint sind (siehe Meinunger, 2008). Somit ist gerechtfertigt, Goethe als standarddeutschen Referenztext einzustufen und Mann, da seine Novelle zeitlich später erschien, als von der Norm abweichend beziehungsweise diese übertreibend zu interpretieren. Es fand also kein Anstieg im Genitivgebrauch statt, bevor ein deutlicher Schwund auftrat.

Abschließend kann ein deutlicher Genitivschwund innerhalb der letzten beiden Jahrhunderte mit Hilfe des erstellten Perlskripts gezeigt werden. Die Ergebnisse machen deutlich, dass der Gebrauch des Genitivs insgesamt von Goethe zu Timm zurückging,

während in der zeitlich dazwischen liegenden Novelle „Der Tod in Venedig“ von Thomas Mann ein deutlich überhöhter Genitivgebrauch auftritt. Die Genitivtypen, die dabei am stärksten in ihrer Verwendung zurückgehen, sind der adnominale und der adverbale Genitiv. Nichtsdestotrotz sind beim adnominalen Genitiv vor allem die adnominalen Genitive mit Possessivpronomen, die s-Genitive und die Relativpronomen in ihrer Verwendung stabil geblieben. Diese drücken zum Großteil ein Besitzverhältnis aus und sind vermutlich gerade aus diesem Grund noch gut erhalten. Zusätzlich lässt sich beim s-Genitiv die Stabilität eindeutig auf die innere Form der Genitive zurückführen. Innerhalb dieser Gruppe verschwindet der vorangestellte Genitiv fast ganz, was auch von der Literatur bestätigt wurde. Der adverbale Genitiv wird im Laufe der Zeit immer weniger benutzt. Die dabei auftretenden Verben scheinen insgesamt nicht mehr verwendet zu werden oder treten in nicht-genitivischen Umgebungen auf. Ansonsten bleiben freie Genitive und Präpositionen, die Genitiv auslösen, zahlenmäßig ebenfalls stabil. Vergleicht man jedoch die einzelnen Vorkommnisse, so entstehen große Zweifel hinsichtlich der Stabilität und diese kann nicht vollständig angenommen werden. Im Gegenteil dazu kann ein Schwund des Genitivs bei den genitivauslösenden Adjektiven zwar prozentual gesehen gezeigt werden, jedoch unterstützen die ausgegebenen Genitive diese Zahlen nicht. Weiterhin ist es fraglich, ob Konjunktionen, Adverbien und Adjektive, die Genitiv enthalten, in die Untersuchung miteinbezogen werden sollten. Über die konkurrierenden Konstruktionen zum Genitiv, die in den Novellen gefunden wurden, können nur sehr aussageschwache Folgerungen hinsichtlich des Genitivschwundes gemacht werden.

7. Fazit

Mit Hilfe der computerlinguistischen Untersuchung konnte gezeigt werden, dass der Rückgang in der Verwendung des Genitivs aus älteren Sprachstufen bis zur heutigen Zeit anhält. Dies wird durch den deutlichen Schwund des gesamten Genitivs innerhalb der letzten beiden Jahrhunderte von Goethe zu Timm belegt. Allerdings konnte kein Verlust des Genitivs von Goethe zu Mann festgestellt werden, denn dort steigt, im Gegensatz zur ursprünglichen Vermutung, der Genitivgebrauch deutlich an. Dies ist wahrscheinlich auf eine unverhältnismäßig häufige Verwendung des Genitivs von Thomas Mann zurückzuführen. Auch eine Deutung des Genitivs als besonderes Stilmittel ist hier möglich.

Durch die genauere Untersuchung der unterschiedlichen Typen wurde herausgearbeitet, dass nicht nur der adverbale Genitiv, sondern auch der adnominale Genitiv in der heutigen Zeit weniger verwendet werden. Nichtsdestotrotz konnte dieser Schwund nicht eindeutig auf die Verdrängung durch konkurrierende Konstruktionen zurückgeführt werden. Darüber hinaus zeigten sich auch einige Genitive, die im zeitlichen Verlauf stabil geblieben sind. Dabei können durch die Betrachtung der inneren Form der einzelnen Typen die Prozentzahlen, die Aussagen über die Stabilität oder gegebenenfalls den Verlust machen, teilweise unterstützt werden. Dies ist aber nicht immer der Fall. Die Untersuchung der einzelnen Typen ist zwar sinnvoll, jedoch sind weitere Untersuchungen notwendig und die Typen sollten jeweils einer eigenen Untersuchung mit vergrößerten Korpora unterzogen werden. Dabei könnte es von Nutzen sein, mehrere Texte innerhalb des gleichen Zeitraums oder auch des gleichen Autors zu untersuchen, um einige Aussagen zu stärken.

Vor allem in Bezug auf die Konkurrenzen besteht hierbei noch Klärungsbedarf, denn eine detailliertere Analyse scheint dort nötig zu sein. Entgegen der Erwartungen konnte kein possessiver Dativ gefunden werden und es werden unterschiedliche Periphrasen gefunden, die nicht unbedingt anstelle eines Genitivs verwendet werden. Auch die Konkurrenzen anderer Genitivtypen können nicht eindeutig in eine verdrängende Position zum Genitiv gerückt werden.

Zudem wäre es möglich, um beispielsweise die Stärke des Rückgangs in kleineren Zeitstufen zu messen und den überhöhten Genitivgebrauch Manns herauszufiltern, eine andere Novelle aus dem gleichen Zeitbereich auszuwählen. Diesbezüglich müsste das Programm gegebenenfalls angepasst werden.

Außerdem wäre interessant zu sehen, wie sich die Ergebnisse verhalten würden, wenn man zum Beispiel die Konjunktionen, Adverbien und Adjektive, die Genitiv enthalten, nicht einbeziehen oder die Relativpronomen weglassen würde. Dazu müssten lediglich die entsprechenden Stellen im Programm auskommentiert werden. Was die Methodik anbelangt, besteht hier ebenfalls noch Freiraum und es kann weitergeforscht werden.

Ein weiterer Punkt, der hinsichtlich der Methoden verändert werden könnte, ist das *Part-of-Speech Tagging*. Der *TreeTagger* ist aufgrund seiner hohen Fehlerquote und der Kodierungsprobleme nicht das idealste Werkzeug zur vorherigen Bearbeitung der untersuchten Texte. Hierfür ließe sich eventuell eine bessere Alternative finden, um aussagekräftigere Ergebnisse zu erzielen. Zudem könnte ein *Part-of-speech Tagger*, der detailliertere morphologische Auskünfte, vor allem über Genus, Kasus und Numerus der Artikel, geben könnte, von Vorteil sein.

Dennoch gelang es mit den verwendeten Mitteln und durch das entwickelte Programm, den Genitivschwund innerhalb der letzten beiden Jahrhunderte zu verdeutlichen und zu belegen. Die Computerlinguistik sollte folglich mehr in die Forschung miteinbezogen werden und bietet neue Möglichkeiten einen Sachbestand von mehreren Seiten zu beleuchten. Die zu Beginn gestellten Ziele wurden zufriedenstellend erfüllt und das Perlskript bildet eine solide Grundlage, den Genitivschwund weiter zu untersuchen.

Den Abschluss dieser Arbeit soll ein Zitat von André Meinunger (2008) bilden, welchem während dieser Untersuchung besondere Beachtung geschenkt wurde. Dieser Grundsatz sollte nie aus den Augen verloren werden.

„Der seriösen Sprachwissenschaft geht es in erster Linie darum herauszufinden, was Sprecher niemals bilden würden, was sie potentiell bilden (können) und wie dies zu beschreiben, zu charakterisieren und formal nachzuvollziehen ist. Es geht nicht darum, vorzuschreiben, was als richtig zu gelten habe und was nicht.“

(Meinunger, 2008: 45)

Literatur- und Quellenverzeichnis

- Aust, H. (2012).** *Novelle*. 5. Auflage. Stuttgart, Weimar: Verlag J.B. Metzler.
- Bayer, J. (2005).** Displaced and misplaced genitives. In: H. Broekhuis et al. (Hrsg.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter.
- Behaghel, O. (1923).** *Deutsche Syntax. Eine geschichtliche Darstellung. Band 1*. Heidelberg: Winter.
- Christiansen, T. et al. (2012).** *Programming Perl*. 4th Edition. Sebastopol, CA: O'Reilly Media.
- Donhauser, K. (1998).** Das Genitivproblem und (k)ein Ende? Anmerkungen zur aktuellen Diskussion um die Ursachen des Genitivschwundes im Deutschen. In: J. O. Askedal (Hrsg.), *Historische germanische und deutsche Syntax. Akten des internationalen Symposiums anlässlich des 100. Geburtstages von Ingerid Dal, Oslo, 27.9-1.10.1995*.
- Duden online (2013).** Letzter Zugriff am 10.03.2013 unter <www.duden.de>.
- Eckermann, J. P. (1981).** *Gespräche mit Goethe in den letzten Jahren seines Lebens*. Frankfurt a. M.: Insel Verlag. Letzter Zugriff am 10.03.2013 unter <<http://gutenberg.spiegel.de/buch/1912/79>>.
- Fischer, A. (1987).** Das Genitivobjekt und damit konkurrierende Objekte nach Verben in Leipziger Frühdrucken. In: J. Schildt (Hrsg.), *Zum Sprachwandel in der deutschen Literatursprache des 16. Jahrhunderts. Studien – Analysen – Probleme. (Bausteine zur Sprachgeschichte des Neuhochdeutschen 63.)* Berlin: Akademie, 267-324.
- Fischer, A. (1992).** Varianten im Objektbereich genitivfähiger Verben in der deutschen Literatursprache (1570-1730). In: J. Schildt (Hrsg.), *Aspekte des Sprachwandels in der deutschen Literatursprache 1530-1730. (Bausteine zur Sprachgeschichte des Neuhochdeutschen 66.)* Berlin: Akademie, 273-324.
- Fleischer, J. & Schallert, O. (2011).** *Historische Syntax des Deutschen: Eine Einführung*. Tübingen: Narr, 83-101.
- Goethe, J.W. von (1828).** *Novelle*. Letzter Zugriff am 10.03.2013 unter <<http://www.gutenberg.org/cache/epub/2320/pg2320.txt>>.
- Hentschel, E. (2010a).** Genitivobjekt. In: E. Hentschel (Hrsg.), *Deutsche Grammatik*. Berlin, New York: De Gruyter, 104.
- Hentschel, E. (2010b).** Genitivreaktion. In: E. Hentschel (Hrsg.), *Deutsche Grammatik*. Berlin, New York: De Gruyter, 104.
- Hielscher, M. (2007).** *Uwe Timm*. München: Deutscher Taschenbuch Verlag.
- Jurafsky, D. & Martin, J. H. (2009).** *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Second Edition. Upper Saddle River, NJ: Pearson Education.
- Kiefer, H. (1910).** *Der Ersatz des adnominalen Genitivs im Deutschen*. Leipzig: August Hoffmann.
- Komodo Editor (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.activestate.com/komodo-edit>>.
- Kurzke, H. (2009).** *Thomas Mann. Ein Portrait für seine Leser*. München: Verlag C.H. Beck.

- Lenz, B. (1996).** *Adverbale Genitive im Deutschen. (Arbeitspapiere des Sonderforschungsbereichs 282 „Theorie des Lexikons“ 77.)* Düsseldorf: Heinrich Heine Universität.
- Lindgren, K. B. (1969).** Diachronische Betrachtungen zur deutschen Satzstruktur. In: H. Moser (Hrsg.), *Sprache. Gegenwart und Geschichte. Probleme der Synchronie und Diachronie. (Sprache der Gegenwart 5.)* Düsseldorf: Schwann, 147-158.
- Lipavic Oštir, A. (2010).** Genitiv. In: E. Hentschel (Hrsg.), *Deutsche Grammatik.* Berlin, New York: De Gruyter, 99-104.
- Mann, T. (1912).** *Der Tod in Venedig.* München: Hyperionverlag Hans von Weber. Letzter Zugriff am 10.03.2013 unter <<http://www.gutenberg.org/cache/epub/12108/pg12108.txt>>.
- Meinunger, A. (2008).** *Sick of Sick? Ein Streifzug durch die Sprache als Antwort auf den Zwiebfisch.* Berlin: Kulturverlag Kadmos.
- Novellendefinition Duden (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.duden.de/rechtschreibung/Novelle>>.
- Penn Treebank (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/Penn-Treebank-Tagset.pdf>>.
- Perl (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.perl.org/get.html>>.
- Plank, F. (1980).** Encoding grammatical relations: Acceptable and unacceptable non-distinctness. In: J. Fisiak (Hrsg.), *Historical Morphology. (Trends in Linguistics 17.)* The Hague: Mouton, 289-325.
- Plank, F. (1995).** Entgrammatisierung – Spiegelbild der Grammatisierung? In: N. Boretzky et al. (Hrsg.), *Natürlichkeitstheorie und Sprachwandel. (Bochum-Essener Beiträge zur Sprachwandelforschung 22.)* Bochum: Brockmeyer, 199-219.
- Rausch, G. (1897).** Zur Geschichte des deutschen Genitivs seit mittelhochdeutscher Zeit. Darmstadt: Otto Hof.
- Schwartz, R. L. et al. (2005).** *Learning Perl.* 4th Edition. Sebastopol, CA: O'Reilly Media.
- Sommerfeldt, K.-E. (1966).** Zu einigen Entwicklungstendenzen im Satzbau der deutschen Sprache. *Deutsch als Fremdsprache* 3, 4. 34-39.
- STTS Table (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html>>.
- STTS Guide (2013).** Letzter Zugriff am 10.03.2013 unter <ftp://ftp.ims.uni-stuttgart.de/pub/corpora/stts_guide.pdf>.
- Timm, U. (2008).** *Die Entdeckung der Currywurst.* 13. Auflage. München: Deutscher Taschenbuch Verlag.
- TreeTagger (2013).** Letzter Zugriff am 10.03.2013 unter <<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>>.

Anhang

Anhang 1: Auszüge aus der Outputdatei *ven-output.txt*

Anhang 2: Auszüge aus der Outputdatei *cur-output.txt*

Anhang 3: Beigefügte CD-ROM, die die Programme *preprocess.pl* und *genitivex.pl*, sowie die Textdateien *goethe_novelle.txt*, *mann_todinvendig.txt* und *timmm_currywurst.txt* enthält. Weiterhin sind die Textdateien, welche die durch *preprocess.pl* bearbeiteten Texte einschließen, *novelle.txt*, *venedig.txt* sowie *currywurst.txt* auf der CD-ROM zu finden. Zudem sind die vom *TreeTagger* annotierten Dateien *nov-tag.txt*, *ven-tag.txt* und *cur-tag.txt* in Anhang 3. Außerdem beinhaltet Anhang 3 die vollständigen Outputdateien *nov-output.txt*, *ven-output.txt* und *cur-output.txt*.

Anhang 1

Zeile 69-76 (Adnominale Genitive mit definitivem Artikel):

Dichter NN des ART Friedrich NE
Jahren NN der ART Unruhe NN
Repliken NN des ART Gesprächs NN
Inferno NN der ART Lazarette NN
Lazarette NN des ART Siebenjährigen NN
Neugier NN der ART Nerven NN
Ende NN des ART Mai NN
Insel NN der ART Adria NE

Zeile 107-137 (Adnominale Genitive mit definitivem Artikel):

Haus NN des ART Aides ADJA
Flüstern NN des ART Gondoliers NN
Gesichter NN der ART Lagune NN
Begegnisse NN des ART Einsam NN
die PDS des ART Geselligen NN
Erscheinungen NN der ART Herreise NN
Gemüt NN des ART Reisenden NN
Terrasse NN der ART Seeseite NN
Teil NN der ART Hotelgäste NN
Erwartung NN des ART Essens NN
Uniform NN der ART Gesittung NN
Spielarten NN des ART Menschlichen NN
Miene NN des ART Amerikaners NN
Vollendung NN der ART Form NN
Gefälligkeit NN der ART Gestalt NN
Bestandteil NN der ART Vornehmheit NN
Regung NN des ART Bedauerns NN
Probleme NN der ART Form NN
Einflüsterungen NN des ART Traumes NN
Geruch NN der ART Lagune NN
Schwere NN der ART Augenlider NN
Klappern NN des ART Teegerätes NN
Hand NN des ART Portiers NN
Eintritt NN des ART Langschläfers NN
Haltung NN des ART Oberkörpers NN
Bewegung NN der ART Kniee NN
Schönheit NN des ART Menschenkindes NN
Charakter NN des ART Anzugs NN
Blüte NN des ART Hauptes NN
Haupt NE des ART Eros NN
Geringel NN des ART Haares NN

Zeile 400-405 (Adnominale Genitive mit definitivem Artikel):

Ehrgeiz NN der ART großen ADJA Hotels NN
Aufsetzen NN des ART weisbeschuhnten ADJA Fuses NN
Bleiben NN des ART trefflichen ADJA Adgio NN
Geschäftsgeist NN der ART gesunkenen ADJA Königin NN
Wände NN des ART schmalen ADJA Kanals NN
Unkenntnis NN der ART eigenen ADJA Wünsche NN

Zeile 590-595 (Adnominale Genitive mit Possessivpronomen):

Gefolgschaft NN seines PPOSAT Liebhabers NN
Überwinder NN seines PPOSAT Wissens NN
Meisterhaltung NN unseres PPOSAT Styls NN
Gebärde NN seiner PPOSAT eigentümlichsten ADJA Figuren NN
derjenigen PDAT seines PPOSAT ersten ADJA Aufenthaltes NN
Beobachter NN ihrer PPOSAT menschlichen ADJA Gemeinschaft NN

Zeile 620-625 (Adnominale Genitive mit Demonstrativpronomen):

Apotheose NN dieses PDAT Befehlswordes NN
Dichter NN all PIAT derer PDS
Folge NN dieser PDAT wiedergeburt NN
Haltung NN dieser PDAT Frau NN
Gleichtakt NN dieses PDAT Daseins NN
Milde NN dieser PDAT Lebensführung NN

Zeile 660-665 (Adnominale Genitive mit Adjektiv ohne Artikel):

Art NN schweifender ADJA Unruhe NN
Zufuhr NN neuen ADJA Blutes NN
Möglichkeit NN sittlicher ADJA Entschlossenheit NN
Merkmale NN fremder ADJA Rasse NN
Inbegriffleitend ADJD tätiger ADJA Tugend NN
Ausführung NN groser ADJA Pläne NN

Zeile 754-779 (Adnominale s-Genitive):

Leben NN Friedrichs NE
Schillers NE Raisonement NN
Friedrichs NE Heldenleben NN
Aschenbachs NE Vorführungen NN
Aschenbachs NE Ansuchen NN
Aschenbachs NE Sinn NN
Landungsplatze NN Venedigs NE
Tadzios NE Zähne NN
Innern NN Venedigs NE
Überzeugung NN Aschenbachs NE
Aschenbachs NE Entschlus NN
Anhauch NN Okeanos NE
Mauern NN Athens NE
Tadzios NE Gegenwart NN
Tadzios NE Schönheit NN
Rosse NN Poseidons NN
Aschenbachs NE Wohnplatz NN
Tadzios NE Auftreten NN
Tadzios NE Augen NN
Gäschen NN Venedigs NE
Innern NN Venedigs NE
Aschenbachs NE Tischchen NN
Aschenbachs NE platz NN
Venedigs NE Obrigkeit NN
Medizinalbeamte NN Venedigs NE
Mengen NN weines NN

Anhang 2:

Zeile 195-200 (Adnominale Genitive mit definitem Artikel):

Streifen NN der ART übereinandergenähten ADJA Felle NN
Schleife NN der ART feinen ADJA Goldkordel NN
eine PIS der ART selbst ADV eingelegten ADJA Gurken NN
Berlin NE der ART späten ADJA fünfziger ADJA Jahre NN
eine PIS der ART fünf CARD echten ADJA Havannas NN
eine PIS der ART kostbaren ADJA englischen ADJA Zigaretten NN

Zeile 551-559 (Freie Genitive):

eines ART Tages NN
eines ART Tages NN
Eines ART Tages NN
Eines ART Tages NN
Eines PIS Nachts ADV
eines ART Tages NN
eines ART Morgens NN
eines ART Tages NN
eines ART Tages NN

Zeile 568-582 (Prä- und Postpositionen mit Genitiv):

während APPR des ART Krieges NN
dazu PAV voller ADJA Lust NN
während APPR eines ART Bombenangriffs NN
, \$, voller ADJA Misstrauen NN
wegen APPR der ART Brandgefahr NN
wegen APPR des ART Currys NN
, \$, voller ADJA Ungeduld NN
dank APPR seines PPOSAT Reiterabzeichens NN
um APPR Gottes NN willen APPO
wegen APPR seiner PPOSAT Nazimitgliedschaft NN
während KOUS des ART Krieges NN
während APPR der ART Fahrt NN
während APPR der ART Schätzung NN
Um APPR Gottes NN willen APPO
wegen APPR seines PPOSAT zerschossenen ADJA Fuses NN

Zeile 591-593 (Präposition mit von oder Dativ anstatt Genitiv):

möglicherweise ADV wegen APPR Fahnenflucht NN
wegen APPR den ART Russen NN
inmitten APPR von APPR angestaubten ADJA Zigarren NN

Anhang 3

CD-ROM:

- preprocess.pl
- genitivex.pl
- goethe_novelle.txt
- mann_todinvendig.txt
- timm_currywurst.txt
- novelle.txt
- venedig.txt
- currywurst.txt
- nov-tag.txt
- ven-tag.txt
- cur-tag.txt
- nov-output.txt
- ven-output.txt
- cur-output.txt