

Informatik erschließt Literatur.

Der *discours citant* in der französischen Romanliteratur
zwischen 1750 und 1920.

Susanne Mocken



Informatik erschließt Literatur.

Der *discours citant* in der französischen
Romanliteratur zwischen 1750 und 1920.

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs-Universität
Freiburg im Breisgau

vorgelegt von

Susanne Mocken
aus Essen

Wintersemester 2012/13

Informatik erschließt Literatur.

Der *discours citant* in der französischen Romanliteratur zwischen 1750 und 1920.

Titel der eingereichten Dissertation:

Der *discours citant* in der französischen Romanliteratur zwischen 1750 und 1920.
Entwicklung und Anwendung eines Software-Tools zum Zwecke einer automatisierten
quantitativen Korpusanalyse.

Susanne Mocken

Erstgutachter

Prof. em. Wolfgang Raible

Romanisches Seminar der
Albert-Ludwigs-Universität Freiburg

Zweitgutachter

Prof. Dr. Stefan Pfänder

Romanisches Seminar der
Albert-Ludwigs-Universität Freiburg

Drittgutachter

Prof. Dr. Christian Mair

Englisches Seminar der
Albert-Ludwigs-Universität Freiburg

Vorsitzender des Promotions-
ausschusses der Gemeinsamen
Kommission der Philologischen,
Philosophischen und Wirtschafts-
und Verhaltenswissenschaftlichen
Fakultät

Prof. Dr. Bernd Kortmann

Datum der Disputation: 19. Februar 2013

Inhaltsverzeichnis

1	Einleitung	9
1.1	Motivation	10
1.2	Methodisches Vorgehen	11
1.3	Vorschau	15
2	Der <i>discours citant</i>: unverzichtbares Begleitmerkmal der Figurenrede	17
2.1	Überlegungen zur Figurenrede im Roman	18
2.2	Eigenschaften des <i>discours citant</i>	26
2.2.1	Bezeichnung, Funktion und Struktur	28
2.2.2	<i>Verba communicandi</i> innerhalb des <i>discours citant</i>	33
2.2.3	Konstruktionstypen innerhalb des redebegleitenden <i>discours citant</i>	35
3	Datengrundlage und Methodik	41
3.1	Das Arbeitskorpus	41
3.1.1	Quellen	41
3.1.2	Qualität der Quellen	42
3.1.3	Verwendete Werke	43
3.2	Für die Untersuchung des <i>discours citant</i> verwendete <i>verba communicandi</i>	45
3.3	Konzeption und Arbeitsweise des Programms DCAnalyse	47
3.3.1	Vor dem Start von DCAnalyse	48
3.3.2	Programmstart	49
3.3.3	Ermitteln der direkten Rede	51
3.3.4	Exkurs: Generieren von Listen mit konjugierten Verbformen	53
3.3.5	Tagging mit der Software TreeTagger	54
3.3.6	Redeeinleitung: Tagging und Ermitteln der <i>verba communicandi</i>	57
3.3.7	Suche nach Redebegleitung und Bestimmen der <i>verba communicandi</i>	61
3.3.8	Ermitteln der verschiedenen Konstruktionstypen	62
3.3.9	Export der Ergebnisse	68

3.4	Hürden beim Programmieren	76
3.4.1	Mangelhafte Textformatierung	76
3.4.2	Durch Tagging induzierte Fehler bei der automatischen Auswertung	78
3.5	Programmende	80
4	Darstellung und Diskussion der Ergebnisse	81
4.1	Allgemeine Eigenschaften der Figurenrede	82
4.1.1	Verhältnis der direkten Rede zum Erzähltext	82
4.1.2	Verteilung Redeeinleitung - Redebegleitung - alleinstehende Dialogstellen	87
4.2	Merkmale des redebegleitenden <i>discours citant</i>	90
4.2.1	Verwendete <i>verba communicandi</i> in der Redebegleitung	90
4.2.2	Verteilung und Entwicklung der Konstruktionstypen	113
4.2.3	Verteilung der VC-Types innerhalb der hochfrequenten Konstruktionstypen	128
4.2.4	Type-Token-Relation der <i>verba communicandi</i>	144
4.2.5	Refunktionalisierung der <i>gérondif</i> -Paraphrase?	156
4.2.6	<i>dire</i> und <i>parler</i> als Teil der <i>gérondif</i> -Paraphrase	172
4.3	Vergleich mit bisherigen Untersuchungen	177
5	Fazit und Ausblick	183
A	Verwendete <i>verba communicandi</i> in der Redebegleitung	191
A.1	Ermittelte VC-Types und Anzahl der VC-Tokens	191
A.1.1	Übersicht VC-Types: Bernardin de Saint-Pierre und Voltaire	191
A.1.2	Übersicht VC-Types: Victor Hugo, George Sand, Stendhal und Eugène Sue	192
A.1.3	Übersicht VC-Types: Honoré de Balzac	195
A.1.4	Übersicht VC-Types: Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert	197
A.1.5	Übersicht VC-Types: Émile Zola	200
A.1.6	Übersicht VC-Types: Henri Alain-Fournier, Marcel Prévost, Marcel Proust und Colette	202
B	Type-Token-Relation der <i>verba communicandi</i>	205
C	Konstruktionstypen innerhalb des redebegleitenden <i>discours citant</i>	207

Danksagung	209
Abbildungsverzeichnis	211
Literaturverzeichnis	215

Kapitel 1

Einleitung

Die Digitalisierung von Dokumenten im großen Stil begann gegen Ende der 1990er Jahre, begünstigt vor allem durch die stetige Verbesserung von Hard- und Software, besonders was Prozessorleistung, Speichermedien, Scanner und Texterkennungssoftware angeht. Seither sind die Bestände von Online-Textdatenbanken kontinuierlich gewachsen, und täglich kommen weltweit neue Digitalisate aus den verschiedensten Bereichen hinzu.

Die Motive für die Erschließung allerlei gedruckten Materials und seine Überführung in elektronische Formate sind vielfältig und reichen vom Anspruch der Bewahrung nationalen Kulturgutes bishin zur Forderung nach einem barrierefreien Zugang zu Informationen jeglicher Art für jedermann. Für Geisteswissenschaftler, besonders aber Linguisten und Literaturwissenschaftler, stellen die verschiedenen Datenbanken eine Fundgrube von unschätzbarem Wert dar, weil nicht nur etliche Dokumente virtuell von überall aus in der Welt rund um die Uhr konsultiert werden können, sondern weil sich durch das Vorhandensein digitaler Fassungen eines Dokuments auch völlig neue Wege für die Analyse von Sprache und Literatur eröffnen.

Während in den Vereinigten Staaten von Amerika bereits ab den 1970er Jahren mit dem *Project Gutenberg*¹ eines der weltweit ersten Digitalisierungsprojekte² startete, lange bevor es das Internet geben sollte, war und ist der Vorreiter in Frankreich bezüglich der Digitalisierung von Bibliotheksbeständen stets das Projekt *Gallica* der Bibliothèque nationale de France (BnF) gewesen. Deren Handschriften und gedruckte Dokumente wurden ab dem Beginn der 1990er Jahre systematisch gescannt³ und sind seit 1997

¹ Vgl. http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart.

² Optische Lesegeräte, d.h. Scanner, gab es zu dieser Zeit noch nicht, weshalb gedruckte Texte stattdessen von Hand abgetippt wurden!

³ Vgl. <http://www.bnf.fr/fr/professionnels/numerisation.html>.

über das Internet zugänglich. Seither ist die Zahl der online abrufbaren Dokumente kontinuierlich angestiegen.⁴ Parallel dazu existierten weitere kleine Projekte, wie z.B. die *Association des Bibliophiles Universels*⁵, die zeitgleich mit *Gallica* ihre Arbeit aufnahm, aber seit 2002 keine neuen Texte mehr zur Verfügung gestellt hat, oder die bis 2010 zur Verfügung gestellte Textsammlung *Foire aux textes*⁶ des *Ministère des affaires étrangères et européennes*.

1.1 Motivation

Mit der Umwandlung eines gedruckten Werkes in ein digitales Format allein ist es jedoch nicht getan, da die Digitalisate erst dann einen echten Mehrwert haben, wenn es Werkzeuge gibt, die sie sinnvoll erschließen helfen. Verglichen mit der Fülle der Digitalisate ist die Anzahl der Software-Tools, die das Arbeiten mit elektronischen Textfassungen ermöglichen, jedoch überschaubar. Dieser Schiefelage entgegenzuwirken ist eines der Ziele, die die Verfasserin mit der vorliegenden Arbeit verfolgt. Sie stellt eine Software zur Verfügung, die eine so weit wie möglich automatisierte Korpusanalyse am Beispiel des *discours citant*, der Redeeinleitung und Redebegleitung in literarischen Prosatexten, durchführt. Hier ist anzumerken, dass bereits das der Auswertung zugrundeliegende Datenmaterial so umfangreich⁷ ist, dass weder das Zusammentragen der Datengrundlage noch die sich daran anschließende Analyse der Fundstellen von einer Einzelperson ohne Rückgriff auf elektronische Hilfsmittel in vertretbarer Zeit zu schaffen wäre.

Der *discours citant* bietet sich als Untersuchungsgegenstand an, weil er bislang erstaunlich selten als solcher innerhalb des Spannungsfeldes zwischen Mündlichkeit und Schriftlichkeit wahrgenommen wurde, was sich vor allem in der nicht gerade zahlreich

⁴ Als Beispiel: Am 30. April 2006 waren in Gallica, der elektronischen Bibliothek der Bibliothèque nationale de France, bereits 90.000 digitalisierte Monographien im Bildformat enthalten, davon rund 1.200 Textdigitalisate (http://www.bnf.fr/fr/professionnels/anx_pol_num/a.Charte_documentaire_de_Gallica_1997_2007.html). Im August 2012 belief sich die Zahl der Monographien bereits auf 372.468, darunter 168.824 Dokumente, die im Textformat vorliegen. Die jeweils aktuellen Bestände kann man auf der Seite http://gallica.bnf.fr/Search?p=1&lang=FR&adva=1&adv=1&t_typedoc=livre&reset=true abfragen: Die Anzahl der Dokumente ist in der Navigationsleiste links unter «*Type de document*», die Zahl der Dokumente im Textformat weiter unten unter «*Mode texte*» zu finden.

⁵ Vgl. <http://abu.cnam.fr/INFO/>.

⁶ http://www.diplomatie.gouv.fr/fr/enjeux-internationaux_830/cooperation-culturelle-medias_1031/collection-textes_5281/index.html. Leider gaben weder die Webseite noch die herunterladbaren Dokumente Aufschluss über den Ursprung der Digitalisate. Aufgrund einer Umgestaltung der Webseite ist der Link inzwischen nicht mehr gültig, und auch die Sammlung selbst ist nicht mehr zugänglich.

⁷ Das Arbeitskorpus umfasst 70 Romane mit ca. 9,8 Mio. Tokens. Eine ausführliche Beschreibung des Korpus erfolgt in Kapitel 3.1.

vorhandenen Sekundärliteratur widerspiegelt. Vor allem quantitative Studien sucht man hier vergeblich, und die wenigen qualitativen Untersuchungen zum Thema geben keinerlei Auskunft darüber, wie die zitierten Zahlenwerte zustandekommen. Der *discours citant* ist jedoch – zumindest, was die vergangenen 200 Jahre angeht – aus der Figurenrede nicht wegzudenken. Schon alleine deswegen ist eine genauere Analyse seiner Funktionsweise und der vollzogenen Entwicklungen anhand einer breiteren Textgrundlage als der bisherigen, die sich auf einzelne Romantexte beschränkt, geboten. Eine weitere Ursache für sein bisheriges Schattendasein, verglichen mit der Erforschung der indirekten und erlebten Rede, liegt sicherlich auch darin, dass das Erzeugen einer soliden Datenbasis ohne weitere Hilfsmittel einen langen Atem erfordert. So bemerkt Vivienne Mylne bezüglich des Wandels der *gérondif*-Konstruktion als adverbialer Ergänzung (z.B. *dit-il en brillant*) des Redeverbs zum eigenständigen Redeverb (z.B. *brailla-t-il*) völlig zutreffend:

«[...] au dix-neuvième siècle les verbes expressifs ont pris le pas sur les verbes résumés et opérationnels ; mais pour s'assurer de ce fait (ou du contraire), il faudrait dépouiller des centaines de romans. Bien que le sujet ne soit pas sans intérêt, il ne vaut peut-être pas le travail fastidieux qu'exigerait une telle étude.»⁸

Dieser Einwand, der auch für alle anderen Aspekte des *discours citant* gilt, ist dank der im Zuge dieser Arbeit entstandenen Software entkräftet, so dass der Analyse der Redebegleitung und auch der Redeeinleitung nichts mehr im Wege steht.

1.2 Methodisches Vorgehen

Das Internet ermöglicht den Zugang zu etlichen Online-Datenbanken, die eine riesige Menge an analysierbarem Material zumeist sogar kostenlos bereitstellen. Für linguistische Analysen ist in der Regel die Durchsicht und Auswertung großer Datenmengen erforderlich. Je größer der Umfang des auszuwertenden Materials, desto mehr Zeit benötigt man für die Sichtung, Aufbereitung und Bearbeitung der Daten. Liegen diese bereits in elektronischer (Text-)Form vor, stellt sich als nächstes die Frage nach der Auswahl der Werkzeuge, die bei der Analyse behilflich sein können.

Beschränkt sich die Analyse auf einzelne Wörter oder Kombinationen aus Wörtern, lässt sich diese prinzipiell auch mit den Bordmitteln eines guten Texteditors erledigen. Aber: Die Verwendung der in den Editor integrierten einfachen Suchfunktion setzt die Eingabe einer oder mehrerer konkreter Wortformen voraus. Für die Suche nach Redeverben hieße dies, dass für jeden Verb-Type nach den konjugierten Verbformen (Tokens)

⁸ Vivienne Mylne (1994), S. 42. Diesem Phänomen widmet sich Kapitel 4.2.5 eingehender.

gesucht werden müsste, woraus sich ein Vielfaches an Wortformen ergibt, die allesamt von Hand in die Suchmaske eingegeben werden müssten. Bei 439 Verb-Types wie im vorliegenden Fall wären dies mehr als 23700 Tokens, die man für die Suche berücksichtigen müsste! Zwar beherrschen viele Texteditoren auch die Suche mittels Regulärer Ausdrücke⁹, mit der man das Problem der Formenvielfalt zumindest teilweise umgehen könnte, doch fehlt dann immer noch die Möglichkeit, die Fundstellen zu konservieren, d.h. in eine Datei zu exportieren. Ohne manuelles „Copy & Paste“ der Suchergebnisse inklusive ihres Kontextes in eine selbst erstellte Datei ist das erneute Betrachten und die Weiterverarbeitung ausgeschlossen. Auch die Möglichkeit zur Stapelverarbeitung mehrerer Dateien ist in der Regel nicht gegeben, so dass für jede Datei dieselben Suchanfragen wiederholt werden müssten. Alles in allem ist dieses Verfahren äußerst umständlich und um so weniger praktikabel, je länger der Text und je größer die Anzahl der zu untersuchenden Texte ist.

Eine weitere denkbare Vorgehensweise wäre das Programmieren spezieller Makros innerhalb von Microsoft Word (mittels VBA) oder OpenOffice / Libre Office Writer (mittels StarBasic). Abgesehen von einem ähnlich hohen Zeitaufwand wie für die von der Verfasserin gewählte Lösung wäre außerdem der Einsatz externer Tools wie TreeTagger deutlich erschwert, wenn nicht sogar ausgeschlossen, und auch der plattformunabhängige Einsatz wäre, zumindest bei Verwendung von Microsoft Word, nicht ohne weiteres gewährleistet.

Ein Teil der in dieser Arbeit vorgestellten Analysen ließe sich unter Umständen mithilfe des Online-Korpus Frantext durchführen, da dieses zum einen bei der Suche den Einsatz Regulärer Ausdrücke erlaubt und zum anderen einige Texte neueren Datums enthält, doch auch die Arbeit mit Frantext ist mit gewissen Einschränkungen verbunden. Da zeitgenössische Texte in der Regel durch Copyright geschützt sind, können zwar Suchanfragen für einen oder mehrere komplette Texte gestellt werden, die Ergebnisanzeige erfolgt aber nur in Auszügen. Der Benutzer hat also keine Möglichkeit, auf den elektronischen Volltext zuzugreifen. Da nur einzelne Texte innerhalb von Frantext kategorisiert, d.h. mit grammatikalischen Meta-Informationen versehen sind, ist der Benutzer bei seiner Wahl des Untersuchungsmaterials eingeschränkt, sobald eine Suche über reine Wortformen nicht mehr ausreicht. Die Benutzung der Datenbank erfolgt außerdem gegen eine jährlich zu entrichtende Gebühr.

⁹ Ein Regulärer Ausdruck (engl. *Regular Expression*) ist eine Zeichenkette, mit der unter Einhaltung bestimmter vorgegebener syntaktischer Regeln nach bestimmten Textmustern gesucht werden kann. Der Ausdruck /*(dé|inter|mé|pré)dire*/ ist ein Suchmuster für die Infinitive *dédire*, *interdire*, *médire* und *prédire*.

Die von der Verfasserin gewählte vierte Methode für die automatisierte Untersuchung der Redeeinleitung und -begleitung, die die meisten der oben genannten Schwierigkeiten umgeht, besteht in der Programmierung einer speziellen Software, DCAnalyse, mit deren Hilfe Text- bzw. HTML-Dateien auf bestimmte, für den *discours citant* typische Merkmale hin untersucht werden. Durch den Einsatz des Programms sind vor allem auch quantitative Analysen, deren Umsetzung noch vor zehn Jahren kaum denkbar schien, möglich, ohne dass das Projekt zum Lebenswerk wird. Zwar bleibt dem Programmbeutzer nicht die Interpretation der automatisch gewonnenen Daten erspart, doch wird durch die Automatisierung möglichst vieler Arbeitsschritte wertvolle Zeit gewonnen, die statt auf die Datenerhebung gleich auf eine qualitative oder eine quantitative Untersuchung verwendet werden kann. Der Arbeitsaufwand für die Entwicklung eigener Software ist zwar hoch, aber gerade auch deswegen vertretbar, weil das fertige Programm für die Analyse weiterer, hier nicht untersuchter Texte eingesetzt werden kann, und außerdem nicht nur für Linguisten, sondern auch für Literaturwissenschaftler von Interesse sein kann. Ein Vergleich der hier gewonnenen Ergebnisse mit denen früherer, in präinformatischer Zeit entstandenen Studien lässt vermuten, dass das manuelle Auswerten prinzipbedingt durch das Exzerpieren entsprechender Textstellen fehleranfällig ist (siehe auch Kap. 4.3). Die Software hingegen hat den großen Vorteil, dass sie ermüdungsfrei ihren Dienst verrichtet und vor allem bei sehr langen Texten anders als das menschliche Auge keine Fundstellen übersieht.

Die Entwicklung der Software erfolgte unter Verwendung der 1987 durch den Linguisten Larry Wall entwickelten Skriptsprache Perl¹⁰, deren besondere Stärke im Verarbeiten großer Textmengen liegt und die mit den Regulären Ausdrücken ein sehr mächtiges Werkzeug zur Datenmanipulation bereitstellt. Die graphische Bedienoberfläche, nachfolgend als GUI¹¹ bezeichnet, wurde mit dem Perl/Tk-Modul erstellt. Die GUI vereinfacht die Bedienung enorm und erspart dem Benutzer den Umgang mit einer Kommandozeile. Einzelne Dateien bzw. ganze Verzeichnisse können komfortabel per Mausklick für die (Stapel-)Verarbeitung ausgewählt werden. Das Programm kann unabhängig von dem für die Entwicklung verwendeten Betriebssystem, d.h. im vorliegenden Fall Linux, plattformübergreifend auf allen gängigen Betriebssystemen eingesetzt werden. Zum Ausführen müssen auf einem Windows-, Mac- oder Linuxsystem lediglich die kostenlose Software Perl, die ebenfalls kostenlose Erweiterung Perl/TK sowie das an der Universität Stuttgart entwickelte Programm TreeTagger¹² installiert sein.

¹⁰ Abk. für engl. *Practical Extraction and Report Language*.

¹¹ Abk. für engl. *Graphical User Interface*.

¹² Details zur Arbeitsweise von TreeTagger finden sich in Kapitel 3.3.5.

DCAnalyse ermöglicht die automatisierte Verarbeitung von elektronischen Text- und HTML-Dateien, wie sie in großer Zahl im Internet frei erhältlich sind. Für die vorliegende Arbeit wurde ein Korpus aus 70 Prosatexten von 16 Autoren mit einer Gesamtwortzahl von 9,8 Millionen Tokens zusammengestellt, das den Zeitraum zwischen 1747 und 1922 abdeckt. Eventuelle Lücken in der Verteilung der Romane ergeben sich vor allem durch das Fehlen elektronischer Textversionen. Es kann auch vorkommen, dass zwar ein E-Text vorliegt, dieser aufgrund seiner Formatierung (z.B. fehlende Akzente) jedoch de facto nicht zu gebrauchen ist.

Jede vom Benutzer ausgewählte Datei wird auf das Vorkommen von direkter Rede hin durchsucht. Bei diesem Vorgang werden alle Fundstellen für die weitere Verarbeitung zwischengespeichert und zusätzlich in einer Exportdatei abgelegt. Anschließend wird weiter unterschieden, ob es sich bei einer Fundstelle um einen Turn mit redegleitendem oder redeeinleitendem *discours citant* handelt, indem geprüft wird, ob der vorhergehende Satz mit einem Doppelpunkt abschließt. Ist dies nicht der Fall, wird angenommen, dass es sich um einen Turn mit Redebegleitung handelt, sofern einer der möglicherweise enthaltenen Teilsätze ein Redeverb¹³ und die für die Redebegleitung übliche Inversion von Subjekt und Verb enthält. Auf diese Weise kann zusätzlich ermittelt werden, wieviele Sätze innerhalb der direkten Rede weder eine Redeeinleitung noch eine Redebegleitung enthalten. Alle Turns werden sortiert nach Redeeinleitung, Redebegleitung und Nichtvorhandensein der ersten beiden in entsprechenden Exportdateien festgehalten. Auch eher allgemeine Eigenschaften der direkten Rede werden automatisiert bestimmt, so z.B. die Gesamtwortzahl, der Anteil der direkten Rede am Gesamttext und die Anzahl der Dialogstellen.

Ausgehend von dieser Vorsortierung in Redeeinleitung und Redebegleitung – alleinstehende Turns werden nur hinsichtlich ihrer Frequenz erfasst – erfolgt die weitere Analyse. Einen wesentlichen Einfluss auf die Qualität der gewonnenen Daten hat die Verwendung des externen Tools TreeTagger¹⁴, mit dessen Hilfe die Turns mit grammatischen Meta-Informationen, sogenannten Tags, versehen werden können. Auf diese Weise bleiben Suchanfragen nicht auf bloße Wortformen und Kombinationen derselben beschränkt, sondern es kann nach Abfolgen von Wortarten gesucht werden, wodurch sich die Suche um einiges flexibler gestaltet.

Wesentliche Merkmale des redeeinleitenden und redegleitenden *discours citant* werden vom Programm automatisch identifiziert, wie etwa die Anzahl der Types und Tokens der verwendeten *verba communicandi*, die Verteilung der Redeverbien in Redeein-

¹³ Eine Übersicht über die überprüften Verben liefert Tabelle 3.2.

¹⁴ Siehe Fußnote 12 auf Seite 13.

leitung und Redebegleitung sowie häufig wiederkehrende Konstruktionen innerhalb der Redebegleitung. Die zahlreichen vom Programm generierten Exportdateien, die nach dem Programmdurchlauf in jedem Internet-Browser konsultiert werden können, runden den Funktionsumfang der Software ab und verschaffen dem Benutzer einen raschen Überblick über die bearbeiteten Texte. Sie können darüber hinaus als Grundlage für zahlreiche weitere linguistische und literaturwissenschaftliche Untersuchungen qualitativer oder quantitativer Natur dienen.

Da die Software unter anderem dazu entwickelt wurde, bisher nicht mögliche quantitative Untersuchungen hinsichtlich des *discours citant* durchführen zu können, und zudem der Umfang des Arbeitskorpus groß gewählt wurde, kann der Fokus für die ab Kapitel 4 vorgestellte exemplarische Analyse der automatisiert erzeugten Datengrundlage nur auf einige wenige Aspekte gelegt werden. In diesem Fall sind dies das allmähliche Entstehen eines diversifizierten Sets aus *verba communicandi* sowie die Entstehung und Entwicklung der einzelnen Konstruktionen innerhalb der Redebegleitung. Die Beschränkung der Auswertung auf den redebegleitenden *discours citant* liegt wiederum in dessen Dominanz gegenüber der Redeeinleitung begründet. Da das Programm auch eine Datengrundlage für die Analyse des redeeinleitenden *discours citant* liefert, steht seiner Untersuchung durch Dritte nichts im Wege.

Bei der Betrachtung der 70 im Arbeitskorpus enthaltenen Texte, was einem Zeitraum von 175 Jahren entspricht, ergibt sich ein differenziertes Bild von der historischen Entwicklung vor allem der Redebegleitung, das den sich allmählich vollziehenden Wandel der Diskurstradition widerspiegelt. Wie man mit relativ einfachen Mitteln selbst ein Werkzeug zur Datengewinnung herstellt, das diese Daten in eine verwertbare Form bringt, auf deren Grundlage weiterführende Analysen erstellt werden können, soll anhand der folgenden Kapitel dargelegt werden.¹⁵

1.3 Vorschau

Kapitel 1 führt in die Bedeutung der Erschließung elektronischer Dokumente ein und stellt die sich daraus ableitende Motivation der Verfasserin sowie die Vorteile der gewählten Methode gegenüber weiteren Bearbeitungsmöglichkeiten vor.

Einen Überblick über das Problem der Figurenrede im Roman sowie eine detaillierte Beschreibung des *discours citant* und seiner kleineren Einheiten, d.h. der *verba*

¹⁵ Allererste Ergebnisse wurden von der Verfasserin unter ihrem Geburtsnamen in Susanne Hoberg (2009) vorgestellt.

communicandi und der häufig wiederkehrenden Konstruktionen, liefert Kapitel 2.

Kapitel 3 informiert über die im Arbeitskorpus enthaltenen Texte und deren Quellen, über die für die Analyse verwendeten *verba communicandi* sowie über den Aufbau und die Arbeitsweise des Programms. Es wird außerdem erläutert, welche Output-Datei welche Ergebnisse enthält, die die Datengrundlage für die in Kapitel 4 behandelten Fragestellungen bilden.

In Kapitel 4 werden die mit Hilfe des Programms DCAnalyse erzeugten Daten vorgestellt, interpretiert und anhand von Tabellen und Abbildungen veranschaulicht. Für die Analyse wurden besonders die im redebegleitenden *discours citant* verwendeten *verba communicandi* sowie deren Verteilung berücksichtigt. Nach einer kurzen Beschreibung der Figurenrede hinsichtlich des Verhältnisses zwischen direkter Rede und dem Erzähltext sowie der Verteilung von Redeeinleitung, Redebegleitung und alleinstehenden Turns werden wichtige Merkmale des redebegleitenden *discours citant* diskutiert. Ein sehr wichtiger Bestandteil der Redebegleitung sind zum einen die darin vorkommenden *verba communicandi*. Eine zusätzliche Möglichkeit zur Beschreibung der *verba communicandi* liegt in der Berechnung ihrer Type-Token-Relation, die Aufschluss darüber gibt, welche Autoren sich um einen variantenreichen Einsatz von *verba communicandi* bemühen. Zum anderen ist es wichtig zu erörtern, wie redebegleitende Sätze strukturiert sind, d.h. welche Konstruktionstypen häufig wiederkehren und welche Redeverben bevorzugt darin eingesetzt werden. Ein weiteres Hauptaugenmerk liegt auf der Erörterung der Frage, inwieweit die redebegleitende Konstruktion mit *gérondif* – z.B. «...», *dit-il en murmurant*. – eine Refunktionalisierung erfahren hat, die dazu führt, dass das Verb aus der *gérondif*-Form ab einem bestimmten Zeitpunkt vor allem als Voll-Redeverb – z.B. «...», *murmura-t-il*. – eingesetzt wird. In diesem Zusammenhang wird außerdem die unterschiedliche Distribution von *dire* und *parler* innerhalb der *gérondif*-Paraphrase untersucht. Die Auswertung der Ergebnisse schließt mit einem Vergleich, der die Resultate früherer Studien den selbstgewonnenen Daten gegenüberstellt und diskutiert.

Im Anhang ab Seite 191 werden einzelne Ergebnisse, d.h. die verwendeten *verba communicandi*, deren Type-Token-Relation und die Häufigkeit der Konstruktionstypen innerhalb der Redebegleitung, tabellarisch dargestellt.

Der Quellcode von DCAnalyse sowie einzelne durch das Programm eingebundene Textdateien befinden sich in der Druckausgabe auf der beiliegenden CD-Rom bzw. können zusammen mit der elektronischen Ausgabe heruntergeladen werden.

Kapitel 2

Der *discours citant*: unverzichtbares Begleitmerkmal der Figurenrede

Der Romancier hat verschiedene Möglichkeiten, um die Äußerungen seiner Figuren abzubilden. Für die Redewiedergabe im Französischen (*discours rapporté*) sind dies vor allem die direkte Rede (*discours direct* (DD)), die indirekte Rede (*discours indirect* (DI)) und die erlebte Rede (*discours indirect libre* (DIL))¹. Eine Äußerung ist grundsätzlich zweigliedrig und umfasst sowohl den Vorgang des Äußerns, den *discours citant*, als auch das Produkt des Äußerns, den *discours cité*.² Im Falle des *discours direct* bleibt die Unabhängigkeit des *discours cité* gegenüber dem *discours citant* auf der Schriftebene durch die Verwendung graphischer Marker gewahrt, während beim *discours indirect* die Autonomie des *discours cité* aufgehoben und dem *discours citant* untergeordnet wird.³

Über die Figurenrede im Roman ist viel nachgedacht und geschrieben worden, sowohl von den Schriftstellern selbst, als auch von literaturwissenschaftlicher und linguistischer Seite.⁴ Dem *discours citant* ist hingegen nur wenig Beachtung geschenkt worden, obwohl er aufgrund seiner metasprachlichen Funktion, dem Kommentieren der Art und Weise, wie die Figuren miteinander sprechen, wesentlich zum Gelingen der Figurenrede beiträgt. Neben dieser Kommentarfunktion hat der *discours citant* noch weitere Aufgaben. Er bettet die Figurenrede so in den Erzähltext ein, dass der durch das Einfügen von direkter Rede erzeugte Bruch⁵ deutlich abgemildert wird. Der narrative Kommentar

¹ Vgl. Dominique Mainueneau (1981), S. 98.

² Vgl. Laurence Rosier (1999), S. 126.

³ Vgl. Dominique Mainueneau (1981), S. 98f.

⁴ Zur Vertiefung sei für das Französische auf die Überblicksdarstellungen von Gillian Lane-Mercier (1989), Françoise Rullier-Theuret (2001) und Sophie Marnette (2005) verwiesen.

⁵ Vergleiche auch Grosse (1972): „Äußerlich kennzeichnen zunächst meist die Anführungszeichen den Beginn und das Ende der wörtlichen Rede im epischen Text. Sie signalisieren sozusagen den Bruch in der Erzählsituation, der dadurch entsteht, dass jetzt neben dem Autor [...] weitere Personen zu Worte

„[...] dient der Einführung von in jeder Hinsicht unterschiedlichen Dialogäußerungen in die Ordnung der Narration, in ihre stilistische, ideologische und moralische Ordnung.“⁶

Über die Ursachen des offenkundigen Desinteresses am *discours citant* lässt sich nur spekulieren. Sei es wegen seiner Formelhaftigkeit, die sich durch begrenzte Gestaltungsmöglichkeiten vor allem in der Redebegleitung äußert⁷, sei es wegen der mit viel Aufwand verbundenen Erhebung von auswertbaren Daten⁸: Der *discours citant* hat sich als alleiniges Forschungsgebiet bislang keiner allzu großen Beliebtheit erfreut. Vielleicht lässt sich zu seiner Funktion auch nicht viel mehr sagen, als bereits in der Sekundärliteratur erwähnt wurde. Dennoch wurden immer nur Teilaspekte behandelt, und auch diese meist nur stichprobenartig, d.h. auf der Grundlage eines einzigen Romans. Diesen Untersuchungen durch eine korpusgestützte Analyse eine größere Basis zu geben und aussagekräftige Zahlen zu liefern, die dabei helfen, weitere Aspekte zu beleuchten und Zusammenhänge herzustellen, ist ein Hauptanliegen der vorliegenden Arbeit.

Im Folgenden wird nun zunächst erläutert, welches die Probleme sind, mit denen sich ein Autor beim Verfassen von Figurenrede konfrontiert sieht und welche Überlegungen Literaturwissenschaftler und Linguisten dazu angestellt haben. Anschließend werden die Eigenschaften des *discours citant* genauer erläutert, unter besonderer Berücksichtigung der darin enthaltenen *verba communicandi* und der häufig darin vorkommenden Konstruktionstypen.

2.1 Überlegungen zur Figurenrede im Roman

Seit vielen Jahrhunderten sehen sich Autoren mit der Aufgabe konfrontiert, die Unterhaltungen ihrer Figuren im Roman abzubilden. Dabei handelt es sich um eine Komponente des Schreibens, die dem einen Schriftsteller leichter von der Feder gehen mag als einem anderen, die aber selbst einen Gustave Flaubert rätseln lässt: «Mais comment faire du dialogue trivial qui soit bien écrit? Il le faut pourtant, il le faut.»⁹

Eine universelle Antwort auf die Frage, wie gesprochene Sprache in der Literatur ab-

kommen, die der Rezipient selbst sprechen hört, bzw. liest [...].“ (S. 654) und Francine Cicurel (1994): «*Le lecteur identifie tout de suite la partie dialoguée car elle introduit une rupture dans le continuum narratif du texte ou sur l'espace de la page.*» (S. 130).

⁶ Głowinski (1974), S. 8.

⁷ Vgl. Głowinski (1974), S. 8.

⁸ Vgl. Zitat von Vivienne Mylne (1994) auf Seite 11.

⁹ Zitiert nach Sylvie Durrer (1994), S. 10. Gerade die Figurenrede bei Gustave Flaubert erfreut sich größter Beliebtheit. Zur Vertiefung seien die Arbeiten von Henry (1961), Ullmann (1964), Claudine Gothot-Mersch (1969), Claudine Gothot-Mersch (1981) und Marie-Thérèse Mathet (1988) empfohlen.

zubilden sei, konnten weder Flaubert noch seine Zeitgenossen geben, und auch heute noch tut man sich schwer damit, eben weil für die Konzeption eines Romandialogs andere Prinzipien gelten als für mündliche Kommunikation, die in ihrer Reinform keinerlei Planung unterliegt. Der Autor übernimmt die Rolle des Regisseurs, denn nur er bestimmt darüber, welche Figur zu welchem Zeitpunkt überhaupt reden darf, und definiert über den Grad an Redegewandtheit, den er ihr zuteil werden lässt, ihre Stellung im Romangefüge.¹⁰ Obwohl die Figurenrede einen Einschnitt innerhalb des Erzählflusses bedeutet und die Autonomie des Erzählers einschränkt, bekommt dieser durch ihren Einsatz Gestaltungsmöglichkeiten, wie er sie nur durch Anwendung seiner eigenen erzählerischen Mittel nicht hätte.¹¹

Die russischen Formalisten machten sich bereits seit den Zwanziger Jahren des vergangenen Jahrhunderts Gedanken über die „Erzählerfunktion in der Komposition von Novelle und Roman“¹².

„Gewöhnlich ist der Novellist bestrebt, mit Hilfe der verschiedenen Verfahren den Eindruck einer unmittelbaren Erzählung, einer Improvisation hervorzurufen. [...] Die schriftliche Kultur zwingt ihn auszuwählen, zu verfestigen, zu bearbeiten, aber um so eher ist er bestrebt, wenigstens die Illusion der freien Improvisation zu bewahren.“¹³

Die Illusion der mündlichen Rede – im Russischen mit *skaz* bezeichnet – erzeugt der Autor, indem er sich an der Wahrnehmung von Personen, die aus einem bestimmten Milieu stammen, orientiert, sich jedoch in letzter Instanz nach der Wahrnehmung eines neutralen Lesers richtet.¹⁴ Eine Gratwanderung zwar, aber für den Autor insofern von Nutzen, als er damit verschiedene Effekte in der sprachlichen Wirkung erzeugen kann.

Darüber hinaus charakterisiert der Autor mittels Figurenrede nicht nur seine Figuren und deren Beziehung zueinander, sondern er gibt durchaus auch etwas von sich selbst preis. Indem er sich inmitten des (fiktiven) Geschehens befindet, ist er entweder Zuhörer oder aktiver Gesprächspartner, wodurch er in die Rolle eines Nachrichtenübermittlers schlüpft. Die Figurenrede zeigt außerdem an, dass der Autor die Aussage einer seiner Figuren für so wichtig erachtet, dass sie als solche Teil der Erzählung wird.¹⁵ Die Romanfigur, die bis zum Ergreifen des Wortes nur als Objekt der Erzählung in Erscheinung tritt, wird innerhalb der Redewiedergabe zum Subjekt der Äußerung:

«L'objet devient sujet, le narré devient narrant pour un temps plus ou moins long. Ce renversement total de la situation narrative, qui rompt la continuation de la narration première

¹⁰ Vgl. Evelyne Cosset (1991), S. 155.

¹¹ Vgl. Heuvel (1978), S. 31. Gleiches bemerkt Głowinski (1974), S. 14.

¹² Vinogradov (1969), S. 169.

¹³ Ejchenbaum (1969), S. 163.

¹⁴ Vgl. Vinogradov (1969), S. 195.

¹⁵ Vgl. Heuvel (1978), S. 29.

[...] a des conséquences qui vont au-delà des domaines de l'information, de la perspective et de la construction.»¹⁶

Der russische Literaturtheoretiker Michail Bachtin setzte sich Ende der 1920er Jahre mit der sozialen Natur der Sprache und den sprachlichen Formen auseinander. Sprache lebt und entwickelt sich durch tatsächlich stattfindende sprachliche Kommunikation und nicht durch ein abstraktes Abbild der Formen einer Sprache.¹⁷ Gesprochene Sprache ist niemals individuell, sondern stets sozialer Natur.¹⁸ Die Form und der Stil von Äußerungen werden von der Kommunikationssituation und den Gesprächspartnern bestimmt.¹⁹ Besonders der Roman lebt davon, dass darin sprachlich aktive Individuen vorkommen, deren Äußerungen nicht nur „wiedergegeben und nachgebildet, sondern ‚künstlerisch abgebildet‘ [...] [werden] durch das ‚Wort des Autors‘.“²⁰ Dieses künstlerische Abbilden wird im Roman durch die „Vermischung zweier sozialer Sprachen“²¹ erreicht, was Bachtin als *Hybridisierung* bezeichnet. Dieses Verfahren bildet zwei „Stimmen“ ab: „abbildendes individuelles Bewusstsein und Wille des Autors sowie individualisiertes Sprachbewusstsein und Wille der abzubildenden Person.“²² Das Zeichnen eines „künstlerische[n] Bild[es] der Sprache“ mittels Hybridisierung entbindet den Romanschriftsteller letztlich von dem Anspruch, die sprachliche Wirklichkeit exakt abzubilden.²³

Mit seiner Dissertation leistet Günther (1928) zwar Pionierarbeit bezüglich der Rededarstellung im Französischen, er beschäftigt sich jedoch mehr mit der psychologischen Basis, der Stilwirkung sowie der Entstehungsgeschichte der direkten Rede, so dass die Beschreibung der Redeeinleitung und der Redebegleitung nur eine marginale Rolle spielt.²⁴

Głowinski (1974) äußert sich unter anderem zu den Funktionen des Dialogs im Roman. Die zitierte Rede ist quasi unverzichtbar, denn „[j]egliches Erzählen ist notwendig zugleich ein Erzählen über das Sprechen der Helden.“²⁵ Allerdings legt der narrative Kontext den Umfang der Figurenrede fest, so dass die Narration den „Bezugspunkt des Dialogs“ bildet²⁶. Während der Dialog durch die jeweilige Situation definiert und zudem

¹⁶ Heuvel (1978), S. 29.

¹⁷ Vgl. Bachtin & Volosinov (1977), S. 137.

¹⁸ Vgl. Bachtin & Volosinov (1977), S. 119.

¹⁹ Vgl. Bachtin & Volosinov (1977), S. 125.

²⁰ Vgl. Bachtin (1979), S. 220.

²¹ Bachtin (1979), S. 244.

²² Bachtin (1979), S. 245.

²³ Vgl. Bachtin (1979), S. 251.

²⁴ Günthers Beobachtungen zur Redeeinleitung und Redebegleitung sind in Kapitel 2.2 aufgeführt. Für einen geschichtlichen Abriss der Entwicklung des *discours rapporté* seit der Antike sei außerdem Kapitel 1 in Laurence Rosier (1999) zur Lektüre empfohlen.

²⁵ Głowinski (1974), S. 1.

²⁶ Głowinski (1974), S. 2.

„formal und stilistisch merkmalshaltig“²⁷ ist, gilt die Narration als weitgehend neutral, was jedoch historisch begründet ist: Der Erzähler wurde lange mit dem Autor gleichgesetzt, der als solcher nicht nur seine literarische, sondern auch seine gesellschaftliche Autorität riskiert hätte, wenn er durch Verwendung von Dialekten oder anderen Sprachvarietäten gegen die literarische Etikette verstoßen hätte.²⁸ Abweichungen von der geltenden Sprachnorm mussten daher in die Figurenrede verlagert werden²⁹, was jedoch auch mit Einschränkungen verbunden war, da „die Rede [...] den Situationen und Verhältnissen, in denen sie zustandekommt, zu entsprechen [hatte] [...]“.³⁰ Der Dialog soll die „sozial fixierten Redeweisen“³¹ imitieren, was dann besonders gelungen ist, „wenn der Autor in den Dialogen über die Literatursprache seiner Zeit hinausgeht, wenn er nach Slang, Argot und Mundart greift.“³²

Anna Wierzbicka (1974) beleuchtet die semantischen Aspekte der direkten und indirekten Rede. Bezüglich der direkten Rede stellt sie fest, dass zurückliegende syntaktische und semantische Analysen zum Thema eines völlig übersehen haben, und zwar den spielerischen und zugleich imaginären Charakter der direkten Rede:

„The person who reports another’s words by quoting them, temporarily assumes the role of that other person, ‘plays his part’, that is to say, imagines himself as the other person and for a moment behaves in accordance with this counter-factual assumption.“³³

Das entspricht genau dem, was ein Romanautor beim Verfassen der Figurenrede fortwährend leistet. Sein ständiges Sich-Hineinversetzen in das Denken und Handeln seiner Charaktere kann also nicht genug gewürdigt werden, vor allem, wenn man bedenkt, dass kaum eines der Gespräche, die er abbildet, tatsächlich jemals stattgefunden hat. Einen Vorteil des direkten Zitierens sieht Wierzbicka auch darin, dass der Zitierende nicht Gefahr läuft, die Bedeutung einer Aussage falsch wiederzugeben, falls er dies in eigenen Worten tut.³⁴

Einen sprechakttheoretischen Ansatz verfolgt Elisabeth Gülich (1978). Sie versteht die Redewiedergabe „als eine kommunikative Handlung [...], deren Gegenstand andere kommunikative Handlungen sind“³⁵. Anhand verschiedener Indikatoren kann der Kommunikationsteilnehmer erkennen, dass die Äußerung eines Sprechers ein Zitat darstellt. Zugrunde gelegt wird ein Modell, dem zufolge Kommunikation auf verschiedenen Ebe-

²⁷ Głowinski (1974), S. 2.

²⁸ Vgl. Głowinski (1974), S. 2.

²⁹ Vgl. Głowinski (1974), S. 3.

³⁰ Głowinski (1974), S. 4.

³¹ Głowinski (1974), S. 5.

³² Głowinski (1974), S. 7.

³³ Wierzbicka (1974), S. 272.

³⁴ Vgl. Wierzbicka (1974), S. 279.

³⁵ Elisabeth Gülich (1978), S. 50.

nen stattfindet, in die Kommunikationsakte eingebettet werden können. So handelt es sich bei der Redewiedergabe um

„eingebettete Kommunikationsakte bzw. Kommunikationsakte auf einer zweiten, ggf. weiteren Ebene [...]. Die eingebettete Kommunikationsebene fungiert jeweils als Redegegenstand auf der übergeordneten Ebene; die Kommunikation auf der zweiten Ebene ist also Redegegenstand auf der ersten Ebene, die Kommunikation auf der dritten Ebene ist Redegegenstand auf der zweiten Ebene und zugleich auf der ersten Ebene. Die übergeordnete Ebene ist jeweils Meta-Ebene zur eingebetteten Ebene, das heißt die erste Ebene der Kommunikation ist Meta-Ebene für die zweite, die zweite für die dritte usw.“³⁶

Obwohl Gülich sich an dem Begriff „Redewiedergabe“³⁷ stört, behält sie ihn als Hyperonym für „alle Formen von Kommunikation auf der zweiten bzw. weiteren Ebene“³⁸ bei. Zu den Indikatoren der Redewiedergabe zählen die Sprecherkennzeichnung, die Adressatenkennzeichnung, Angaben zur Kommunikationssituation, ein sprechaktkennzeichnendes Verb, ein sprechaktkennzeichnendes Nomen sowie die Konstitution eines Redegegenstands³⁹. Die Sprecherkennzeichnung erfolgt dadurch, dass der Sprecher als Subjekt eines Verbs in Erscheinung tritt, das einen Sprechakt charakterisiert⁴⁰. Sie kann aber auch wegfallen, obwohl das seltener geschieht als die Auslassung des Indikators „Adressatenkennzeichnung“. Diese muss erfolgen bei Verben, bei denen sie als direktes Objekt fungiert; bei *verba communicandi*, bei denen der Adressat durch ein direktes Objekts angegeben wird, muss sie jedoch nicht stehen. Ebenso fakultativ sind Angaben zur Kommunikationssituation oder zum Redegegenstand, und auch das sprechaktkennzeichnende Verb/Nomen kann ausbleiben⁴¹. Das Fehlen eines oder mehrerer dieser Indikatoren führt nicht zwingend zum Scheitern des Sprechakts der Redewiedergabe, da oft eine Minimalkennzeichnung ausreicht und vermeintliche Lücken auf der Satzebene durch den Kontext geschlossen werden⁴². Jacqueline Authier (1978) geht in ihrer Beschreibung der Redewiedergabe unter anderem der Frage nach, in welchem Maße es sich bei der direkten Rede um eine wortgetreue Wiedergabe des Gesprochenen han-

³⁶ Elisabeth Gülich (1978), S. 53.

³⁷ Zwölf Jahre später ergänzt Elisabeth Gülich (1990), S. 89: „Der Terminus ‚Redewiedergabe‘ scheint mir vor allem deshalb nicht adäquat, weil es sich in einem literarischen Text streng genommen nicht um die ‚Wiedergabe von Rede‘ handelt. Diese Bezeichnung suggeriert, dass die Gespräche irgendwann einmal stattgefunden haben und nun ‚wiedergegeben‘ werden. Man kann sich aber bei einem fiktionalen Text nicht auf eine ‚Originaläußerung‘ beziehen, wie etwa in Zeitungen oder Interviews, wo es tatsächlich um ein Erwähnen oder Wiedergeben aus früheren Kommunikationssituationen geht [...]“. Sie bevorzugt daher den Begriff „erzähltes Gespräch“. Zur Problematik des Begriffs „Redewiedergabe“ siehe auch der Beitrag von Thierry Gallèpe (2002). Die diachronische Entwicklung im Französischen beschreibt wiederum Georgette Stefani-Meyer (2003) ausführlich.

³⁸ Elisabeth Gülich (1978), S. 54.

³⁹ Vgl. Elisabeth Gülich (1978), S. 59.

⁴⁰ Vgl. Elisabeth Gülich (1978), S. 61.

⁴¹ Vgl. Elisabeth Gülich (1978), S. 62.

⁴² Vgl. Elisabeth Gülich (1978), S. 72ff.

delt. An sich signalisiert bereits die Erscheinungsform der direkten Rede die vermeintlich exakte Reproduktion des Gesagten. Dennoch ist die Qualität der Redewiedergabe abhängig von der Sorgfalt, die die wiedergebende Person darauf verwendet:

«Le DD, isolant le message rapporté, peut certes se prêter à l'imitation : il y a loin cependant de cette possibilité, à une propriété constitutive du DD. De fait, l'imitation du ton, du débit, de la voix, de l'accent ou de toute particularité articulatoire, traduit une intention particulière de L et ne correspond pas au cas général.»⁴³

So ist etwa denkbar, dass ein Französisch-Muttersprachler die Aussage eines ausländischen Sprechers möglicherweise zitiert, ohne Bezug auf dessen deutlich hörbaren Akzent zu nehmen, wie etwa in folgendem Beispiel:

(1) *John a dit : «Il fait beau. On va se baigner.»*⁴⁴

Umgekehrt wäre eine Aussage denkbar, deren Redebegleitung explizit auf die Art und Weise der Äußerung hinweist:

(2) *Pour nous faire rire il disait : «C'est une histoire invraisemblable ; je ne veux pas la croire.» avec successivement l'accent bourguignon, alsacien, marseillais et valaisan.*⁴⁵

Die Wiedergabe mündlicher Aussagen umfasst gleichermaßen Dekodierung und Neukodierung, bei welcher der Wiedergebende gar nicht anders kann, als die Phoneme innerhalb der ursprünglichen Lautkette durch seine eigenen Allophone zu ersetzen.⁴⁶ Allein aus diesem Grund kann es sich bei der direkten Rede nicht um eine perfekte Imitation des Gesagten handeln. Auf der Textebene beschränkt sich die Gestaltungsmöglichkeit der wiedergebenden Person auf das, was bis zum Beginn der Anführungszeichen steht. Nach diesem nicht nur optischen Einschnitt nimmt sie sich zurück⁴⁷ und reduziert ihre Aktivität auf die bloße Wiederholung des Gesagten. Durch diese Zurücknahme wird zwar der Eindruck der Unabhängigkeit bzw. Objektivität der direkten Rede verstärkt, was aber keinesfalls bedeutet, dass deren Inhalt ausschließlich die Wahrheit abbildet.⁴⁸ Die Illusion der Unmittelbarkeit der direkten Rede währt um so länger, je mehr sich die Unter-

⁴³ Jacqueline Authier (1978), S. 49. Mit «DD» ist «discours direct», d.h. die direkte Rede, und mit «L» derjenige gemeint, der etwas mittels direkter Rede wiedergibt. Zu den Erkenntnissen Authiers siehe auch Combettes (1989), der kritisiert, dass die Redewiedergabe – dazu zählt er den «discours direct» (DD), den «discours indirect» (DI) und den «discours indirect libre» (DIL) – häufig nur im grammatischen Kontext betrachtet und der Schwerpunkt vor allem auf das hierarchische Verhältnis (S. 111) der einzelnen Komponenten zueinander gelegt werde, wonach ein DD in einen DI und ein DI in einen DIL umgewandelt werden kann. Dieser traditionellen Sichtweise stellt Combettes (1989) die Studien Bachtins sowie Ann Banfields und Jacqueline Authiers gegenüber.

⁴⁴ Jacqueline Authier (1978), S. 49.

⁴⁵ Jacqueline Authier (1978), S. 49.

⁴⁶ Vgl. Jacqueline Authier (1978), S. 50.

⁴⁷ Zur selben Erkenntnis kommt auch Heuvel (1978) auf Seite 24.

⁴⁸ Vgl. Jacqueline Authier (1978), S. 51.

haltung in die Länge zieht.⁴⁹ Dennoch bleibt der Wiedergebende stets der sinnstiftende Textgestalter:

«[...] qu'il le veuille ou non, il tire toutes les ficelles de la situation d'énonciation qu'il rapporte et dans laquelle l'énoncé s'inscrit et prend son sens.»⁵⁰

Deshalb ist auch nicht ausgeschlossen, dass der Wiedergebende eine an sich wahre Aussage manipuliert, indem er sie in einen Kontext rückt, der ihren Sinn verändert oder gar ins Gegenteil verkehrt. Daraus ergibt sich, dass

«[...] le DD, en tant que forme linguistique, manifeste [...] par le clivage que le rapport instaure entre les mots et le cadre de leur production, l'irréductibilité du sens au signifié.»⁵¹

Dominique Maingueneau (1986) stellt ihrerseits in Frage, ob der Terminus *discours <rapporté>* für fingierte Mündlichkeit überhaupt zutreffend ist, da dieser ja nur unter der Voraussetzung existiere, dass man die Umstände der erzählerischen Illusion akzeptiert:

«La narration ne rapporte pas des propos antérieurs qu'elle altérerait plus ou moins, elle les crée de toutes pièces, au même titre que ceux du discours citant. Dans ces conditions, la <fidélité> du discours direct apparaît comme pure convention littéraire.»⁵²

Da es sich bereits beim Erzähltext⁵³ um reine Fiktion handelt, kann man den Wahrheitsgehalt der Figurenrede ebenfalls nicht anzweifeln.⁵⁴

Laut Goetsch (1985) ist „Mündlichkeit in geschriebenen Texten [...] nie mehr als sie selbst, sondern stets fingiert und damit eine Komponente des Schreibstils und oft auch der bewussten Schreibstrategie des jeweiligen Autors.“⁵⁵ Die Figurenrede im Roman ist lediglich ein Konstrukt, das die stilistischen Bemühungen des Autors erkennen lässt:

«Même quand les personnages s'expriment d'une façon qui ressemble au langage parlé plutôt qu'à un style littéraire, ce n'est que fictivement que ces personnages 'parlent' : tous les mots d'un roman sont des mots écrits.»⁵⁶

Es geht weniger darum, „echte“ Mündlichkeit abzubilden, auch weil „Wirklichkeitstreue in der Literaturgeschichte nur in wenigen, allerdings immer wiederkehrenden Strömun-

⁴⁹ Vgl. Vogt (1998), S. 151.

⁵⁰ Jacqueline Authier (1978), S. 53. Siehe auch Marie-Hélène Pérennec (2002): „Hinter dem Wechselspiel der Stimmen zwischen Erzähler und Figur verbirgt sich eine dritte Instanz, der Autor, der zur wirklichen Welt gehört und der letzten Endes die Fäden in der Hand hat.“ (S. 50).

⁵¹ Jacqueline Authier (1978), S. 53. Auf diese Problematik weist auch Bachtin (1979) hin: „Mit Hilfe der Verfahren der Einrahmung kann man eine genau wiedergegebene fremde Äußerung wesentlich verändern. Der gewissenlose und geschickte Polemiker weiß genau, welchen dialogisierenden Hintergrund er den exakt zitierten Wörtern seines Gegners geben muss, um ihren Sinn zu entstellen. [...] Deswegen darf man bei der Untersuchung der Formen der Wiedergabe fremder Rede nicht die Verfahren der Formgebung der fremden selbst von den Verfahren ihrer kontextuellen (dialogisierenden) Einrahmung trennen – das eine ist untrennbar mit dem anderen verknüpft.“ (S. 227).

⁵² Dominique Maingueneau (1986), S. 88.

⁵³ Diesen bezeichnet Dominique Maingueneau (1986) als «discours cité» (S. 85).

⁵⁴ Vgl. Dominique Maingueneau (1986), S. 88.

⁵⁵ Goetsch (1985), S. 202.

⁵⁶ Vivienne Mylne (1994), S. 9.

gen als Darstellungsprinzip“⁵⁷ gilt, sondern vielmehr um eine Imitation derselben, die sich nicht an der (oralen) Realität orientiert, sondern an dem, was die schriftliche Tradition vorgibt⁵⁸. Aufgrund seiner schriftlichen Konzeption und seiner Einbettung in eine Erzählung, innerhalb derer er nur eine von mehreren Textkomponenten darstellt, ist der Romandialog unter den vielen Spielarten fingierter Mündlichkeit (philosophischer Dialog, Filmdialoge, Werbespots, Dramendialoge etc.) derjenige, der von der natürlichen Mündlichkeit am weitesten entfernt ist.⁵⁹

„Echte“ und fiktive Mündlichkeit scheinen auf den ersten Blick einander nicht besonders unähnlich zu sein. Dennoch weisen sie erhebliche Unterschiede auf, die zum einen in der Behandlung der sprachlichen und nicht-sprachlichen Elemente (Prosodie, Akzent, Intonation, Gestik, Mimik), zum anderen in der Verarbeitung der rein sprachlichen Komponente liegen. Während in „echter“ Mündlichkeit alle sprachlichen und nicht-sprachlichen Elemente miteinander verknüpft sind und simultan ablaufen, spielt sich die fiktive Rede allein auf der sprachlichen (Schrift-)Ebene ab. Der Autor kann zwar Auskunft geben über Betonung, Akzent, Tonfall sowie die Figurenrede begleitende Gestik und Mimik, doch trotz verschiedener Gestaltungsmöglichkeiten kann dies für ein schriftliches Medium nur getrennt voneinander, also linear ablaufen. Redewiedergabe, egal ob deren Basis „echt“ oder fiktiv ist, ist außerdem selten vollendet:

„Ob gewollt oder nicht, Redewiedergabe ist immer verbunden mit einem Defizit an Informationen. Schon aus Zeitgründen ist es meist unmöglich, dem Hörer alle Einzelheiten des ursprünglichen Gesprächs mitzuteilen.“⁶⁰

Körperliche Ausdrucksweisen sind ein „eigenes Kommunikationsmittel, das gelernt wird, das wie die Verbalsprache ein Instrument der sozialen Verständigung, also der Interaktion, ist [...]“⁶¹.

„Intonation, Mimik, Gestik und Körperhaltung gewinnen, da im schriftlichen Medium als kommunikative Signale für Verstehensprozesse nicht nötig, in der Feder des Schriftstellers einen eigenen, einen nunmehr bewussten Stellenwert. Der Schriftsteller gestaltet seine erzählten Kommunikationssituationen willentlich komplex, indem er Sprechen und Körper seiner Personen beachtet und mitteilt.“⁶²

Um zu ergründen, wie Körpersprache in vergangenen Jahrhunderten eingesetzt wurde, ist neben dem Sprachsystem und den Kunstformen vor allem die Literatur zu berücksichtigen⁶³, in der sich das soziokulturelle Verständnis der Autoren und ihrer Epoche

⁵⁷ Anne Betten (1994), S. 521.

⁵⁸ Vgl. Haig (1986), S. 2.

⁵⁹ Vgl. Catherine Kerbrat-Orecchioni (1996), S. 332.

⁶⁰ Marie-Hélène Pérennec (2002), S. 42.

⁶¹ Kalverkämper (1997), S. 216.

⁶² Kalverkämper (1991), S. 330.

⁶³ Vgl. Kalverkämper (1997), S. 225.

widerspiegelt.⁶⁴ Besonders die während des Realismus und des sich daran anschließenden Naturalismus entstandenen Romane und Erzählungen geben Aufschluss über das „körpersprachliche[...] Verhalten der Handlungsbeteiligten“⁶⁵.

Eine zusätzliche Aufwertung erfährt die Figurenrede zufolge durch das veränderte Seitenlayout von Druckerzeugnissen, das die Figurenrede optisch vom Rest der Erzählung abhebt. Diese Neuerung beginnt sich ab dem Ende des 18. Jahrhunderts durchzusetzen und ist mit einzelnen Ausnahmen bis in die heutige Zeit die gängige Praxis.⁶⁶

Ab etwa 1820 vollzieht sich innerhalb der Figurenrede ein grundlegender Wandel bezüglich der Darstellung sprechender Romanfiguren:

«[...] on ne s'exprime pas avec les mots seuls, mais avec le corps tout entier. Il ne [...] suffit pas [aux romanciers] d'inventer des personnages, de les faire vivre et penser, ils les font parler, interrompant le récit pur pour y introduire [...] de l'oral dans de l'écrit – qu'est le dialogue à deux ou plusieurs voix ; et les faisant parler, ils ne négligent pas ce qui accompagne la parole, parfois s'y substitue et, de toute façon, la modifie : la gestualité sous toutes ses formes, sans oublier les intonations changeantes du parler et la musicalité de son support vocal.»⁶⁷

Bei verbaler Kommunikation spielt nicht nur eine Rolle, was der Gesprächspartner hört, sondern auch, was er visuell wahrnimmt und ob möglicherweise auch sein Tast- oder Geruchssinn angesprochen wird.⁶⁸

Ein weiteres Indiz für eine Veränderung der Redegestaltung in Prosatexten ist die stetige Zunahme der für den *discours direct* verwendeten *verba communicandi*, die mehr und mehr nicht nur Eigenheiten der Stimme abbilden, sondern auch Aufschluss über das Körperverhalten der sprechenden Figuren geben. Diese Veränderung datiert Vivienne Mylne (1994) auf etwa denselben Zeitpunkt wie Rousset (1998), wenngleich sie erste Anzeichen dafür bereits im 17. Jahrhundert zu erkennen glaubt.⁶⁹

2.2 Eigenschaften des *discours citant*

Der *discours citant*, d.h. die Redeeinleitung und die Redebegleitung verraten dem Leser mehr über die Begleitumstände der direkten Rede und gewähren vor allem eine bessere Innensicht der sprechenden Figuren. Die optische Trennung des *discours citant* vom *discours cité* durch Interpunktion und andere graphische Marker (z.B. Zeilenwech-

⁶⁴ Vgl. Kalverkämper (1997), S. 224.

⁶⁵ Kalverkämper (1991), S. 332.

⁶⁶ Vgl. Vivienne Mylne (1994), S. 189.

⁶⁷ Rousset (1998), S. 78.

⁶⁸ Vgl. Rousset (1998), S.78.

⁶⁹ Vgl. Vivienne Mylne (1994), S. 42.

sel oder Leerzeile) bedeutet nicht nur für den Leser eine Orientierungshilfe⁷⁰, sondern ist darüber hinaus gerade für die automatisierte Verarbeitung von Passagen, die Figurenrede enthalten, unerlässlich.

Der *discours citant* ist innerhalb der Sekundärliteratur meist nur als Begleiterscheinung des *discours cité* bzw. des *discours direct* wahrgenommen und erwähnt worden, und auch seine Benennung ist – im Französischen wie im Deutschen – alles andere als einheitlich. So findet sich im Deutschen unter anderem die Bezeichnung *inquit-Formel*⁷¹ und „narrative(r) Begleitung / Kommentar“⁷² sowie „Redekennzeichnung“⁷³ und „Redeeinbettung“⁷⁴, während im Französischen vor allem *syntagme introducteur*, *phrase introductive*, *formule initiale*⁷⁵ oder *formule/terme de présentation*⁷⁶ als Termini für die Redeeinleitung sowie *incise (déclarative)*, *insertion incidente*⁷⁷ oder *formule en incise*⁷⁸ für die Redebegleitung zu finden sind. Als Hyperonym für beides trifft man im Französischen vor allem auf *attribution du dit*⁷⁹, *discours attributif*⁸⁰, *discours citant*⁸¹, *discours distributif*⁸² und *indices d'attribution*⁸³.

Dass sich der redebegleitende *discours citant* stets stärker im Fokus von Linguisten befand, dürfte zum einen damit zusammenhängen, dass Redebegleitungen sehr viel häufiger vorkommen als Redeeinleitungen und damit das auszuwertende Material umfangreicher ist⁸⁴, und zudem die Besonderheit der Inversion von Subjekt und Verb, die in der Redeeinleitung so nicht existiert, das Interesse an der Beschäftigung damit weckt. Die Beschäftigung mit den *verba communicandi* ist und bleibt jedoch das am häufigsten

⁷⁰ Laurence Rosier (1999), S. 133.

⁷¹ Vgl. Vogt (1998), S. 150.

⁷² Głowinski (1974), S. 8. Der Beitrag wurde aus dem Polnischen ins Deutsche übersetzt, weshalb die Termini vom Übersetzer stammen.

⁷³ Kurz (1966), S. 80. Darunter fällt sowohl die Redeeinleitung als auch die Redebegleitung.

⁷⁴ Müller (1981), S. 136. Die Redeeinbettung umfasst nach Müller sowohl die Redeeinleitung als auch die Redebegleitung.

⁷⁵ Annie Delaveau (1988), S. 125.

⁷⁶ Vgl. Dagmar Almenberg (1958), S. 201.

⁷⁷ Dessaintes (1960), S. 19.

⁷⁸ Annie Delaveau (1988), S. 125.

⁷⁹ Laurence Rosier (1999), S. 131.

⁸⁰ Prince (1978), S. 305.

⁸¹ Dominique Maingueneau (1981), S. 98.

⁸² Halina Suwala (1987), S. 53.

⁸³ Vivienne Mylne (1994), S. 17. Für die Redeeinleitung und die Redebegleitung bevorzugt sie den Terminus *indices d'attribution*, weil *incise* «[...] ne se réfère, littéralement, qu'à des attributions intercalées et ne se rapporte donc pas aux mots qui remplissent cette fonction au commencement ou à la fin d'un énoncé. De plus, le romancier peut attribuer tel énoncé à tel personnage en fournissant un nom, mais sans accompagner ce nom d'un verbe déclaratif».

⁸⁴ Vgl. hierzu Abbildung 4.4.

bearbeitete Gebiet innerhalb des *discours citant*⁸⁵.

2.2.1 Bezeichnung, Funktion und Struktur

Die terminologische Vielfalt deutet bereits an, dass nicht immer eindeutig nach redееinleitendem bzw. redēbegleitendem *discours citant* unterschieden wird. Dennoch wird im Folgenden soweit möglich nach Redeeinleitung, Redeeinleitung und unbegleiteter Rede unterschieden.

Die Redeeinleitung – der deutsche Begriff zeigt dies bereits an – steht immer vor der direkten Rede, die die Leerstelle des direkten Objekts füllt, und schließt mit einem Doppelpunkt.⁸⁶

(158) *Il tressaillit quand une voix douce lui dit tout près de l'oreille :*

(159) – *Que voulez-vous ici, mon enfant ?*⁸⁷

Nach Günther (1928) sind die Hauptformen der Redeeinführung – darunter fällt für den Autor auch die Redebegleitung – „1. Ankündigung mit der Zeile endend, 2. in der Zeile endend, 3. eingeschoben“ genannt, außerdem die nicht so häufig auftretende „freie Ankündigung“ sowie die „uneingeführte Rede“⁸⁸.

Zwei weitere recht früh entstandene stilistische Analysen sind der Beitrag von Gougenheim (1938) und die Arbeit von Dagmar Almenberg (1958), die Darstellung der Figurenrede anhand ausgewählter Romane beschreiben und vergleichen. Gougenheim (1938) stellt anhand einer Analyse von *La Princesse de Clèves* (1678) von Mme de La Fayette und *Dominique* (1863) von Eugène Fromentin fest, dass dies entweder durch einen Einleitungssatz vor der Äußerung oder durch eine Redebegleitung in Mittel- oder Endstellung bewerkstelligt. Darüber hinaus kommt er zu dem Schluss, dass bei beiden Werken trotz ihrer weit auseinanderliegenden Entstehungszeitpunkte ein ähnlicher Stil bezüglich der Redeeinleitung und Redebegleitung vorliegt.⁸⁹

Dagmar Almenberg (1958) bemerkt bezüglich des Gegenwartsromans jedoch:

⁸⁵ Siehe auch Kapitel 2.2.2.

⁸⁶ Vgl. Riegel *et al.* (2009), S. 1011.

⁸⁷ Beide Beispiele, (158) und (159), stammen aus Stendhal, *Le rouge et le noir* (1830). Die Ziffern in Klammern zu Beginn des Beispiels weisen auf die Position des Turns innerhalb des Textes hin. Das Programm DCAnalyse liest jede Datei absatzweise ein und versieht dabei jeden Absatz mit einer eindeutigen Nummer. Auf diese Weise können alle hier zitierten Beispielsätze innerhalb der Output-Dateien leicht wiedergefunden werden.

⁸⁸ Günther (1928), S. 24. Seiner Analyse zur Verteilung der einzelnen Typen liegen vornehmlich Texte des Mittelalters zugrunde; so liefert Günther konkrete Zahlen für die Chansons de geste *Couronnement de Louis* und das *Rolandlied* sowie für einzelne Werke von Chrétien de Troyes, Marie de France und Froissart. Neuzeitliche Prosatexte werden hingegen nicht berücksichtigt.

⁸⁹ Vgl. Gougenheim (1938), S. 320.

„Dans le roman contemporain, les différences dans la manière de présenter les répliques sont assez marquées et assez constantes pour qu'on puisse parler d'un véritable système de présentation, individuel pour chaque auteur“⁹⁰.

Die Autorin unterscheidet für die Redeeinleitung grundsätzlich zwischen einer *formule de présentation* und einer *phrase d'indication*. Erstere unterteilt sie in vier Untergruppen: (a) *terme simple*, (b) *phrase tronquée*, (c) *terme composé* bzw. *périphrase* und (d) *absence totale*⁹¹. Der Terminus (a) steht für einen redeeinleitenden Satz, der ein Redeverb oder eine äquivalente verbale Konstruktion enthält, während (b) diejenigen Einleitungssätze umfasst, die kein Redeverb enthalten, sondern stattdessen entweder mit einer Konjunktion oder auch einem Adverb enden, z.B.

(a) *Jacques raconta :*

– *J'ai dormi comme un voleur sur le quai, sous une bâche ! Et toi ?*⁹²

(b) *Marilhat réfléchit un instant. Puis :*

– *Je ne doute pas de votre habileté.*⁹³

Unter (c) fallen alle redebegleitenden und redeeinleitenden Sätze, die aus einer Kombination aus Redeverb und fester Wendung oder einer Periphrase gebildet werden:

(c1) *M. Paul vint dire au patron :*

– *C'est le monsieur de l'autre fois.*⁹⁴

(c2) *Antoine attendait un «mais», qui vint, en effet, après une brève hésitation :*⁹⁵

Aufgrund ihrer Vielfalt lassen sich diese Periphrasen kaum zu Gruppen zusammenfassen und können daher als „[...] forme la plus personnelle, celle qui porte le plus «la griffe» de l'auteur“⁹⁶ angesehen werden.

Eine *phrase d'indication* weist zwar hinsichtlich Rhythmus und Satzbau Ähnlichkeiten zur *formule de présentation* auf, enthält jedoch kein Redeverb und präzisiert vor allem eine Bewegung, eine Geste, einen Gesichtsausdruck oder die Art und Weise zu reden, z.B.:

Daniel fit un geste vague :

– *Sur un banc là-bas... Et puis, surtout, j'ai erré.*⁹⁷

⁹⁰ Dagmar Almenberg (1958), S. 200.

⁹¹ Auf die *absence totale* geht die Autorin leider überhaupt nicht weiter ein, d.h. sie nennt auch kein Beispiel.

⁹² Dagmar Almenberg (1958), S. 202.

⁹³ Dagmar Almenberg (1958), S. 207.

⁹⁴ Dagmar Almenberg (1958), S. 201.

⁹⁵ Dagmar Almenberg (1958), S. 209.

⁹⁶ Dagmar Almenberg (1958), S. 209.

⁹⁷ Dagmar Almenberg (1958), S. 202.

Laut Rey-Debove & Rey (1993) ist der Terminus *incise* als Bezeichnung für einen redebegleitenden Satz erstmalig 1771 belegt, und zwar im *Dictionnaire de Trévoux*. Eine *incise* ist demnach eine

„proposition généralement courte, tantôt insérée dans le corps de la phrase, tantôt rejetée à la fin, pour indiquer qu'on rapporte les paroles de qqn ou pour exprimer une sorte de parenthèse.“⁹⁸

Bis zum 17. Jahrhundert war innerhalb der *incise* die Verwendung des Pronomens *ce* üblich, analog gebildet zu lat. *Hoc illud dixit*⁹⁹, das an erster Stelle des Einschubs stand und auf die direkte Rede Bezug nahm:

– *Je devois, CE dis-tu, te donner quelque avis.*¹⁰⁰

Die Inversion, die auch nach dem Wegfall des Pronomens erhalten blieb¹⁰¹, erzeugt eine syntaktische Abhängigkeit vom übergeordneten Satz. Dieser Autonomieverlust hat zur Folge, dass ein Einschub nie an erster Stelle eines Satzes oder nach einem Punkt vorkommen kann, sondern stets nur nach einem Komma, dem ein Aussagesatz vorausgeht.¹⁰² Verständlich wird der eingeschobene Satz erst in Verbindung mit dem Aussagesatz.¹⁰³ Trotz seiner Einbettung in die direkte Rede bzw. trotz seiner Position am Ende der direkten Rede wird ein Einschub nicht durch einen syntaktischen Marker untergeordnet.¹⁰⁴

Dessaintes (1960) bezeichnet die Einbettung einzelner Wörter oder auch Wortgruppen in einen Satz bzw. in eine Äußerung¹⁰⁵ als *insertions incidentes* bzw. *incidentes*. Der Einschub unterbricht den Satz nicht nur durch äußere Merkmale wie Kommata oder Gedankenstriche, sondern er verändert auch die Satzmelodie. Er gestattet dem Sprechenden oder Schreibenden, zum Gesagten Stellung zu beziehen und seine subjektive Sichtweise in die Äußerung einfließen zu lassen.¹⁰⁶ Hauptmerkmal einer *incidente* im Sinne der u.a. hier vorgestellten Redebegleitung ist die Änderung der Abfolge Subjekt-Verb in Verb-Subjekt, die Inversion. Diese impliziert eine syntaktische Abhängigkeit des Einschubs vom übergeordneten Satz und bewirkt, dass dieser nur in Verbindung mit diesem einen Sinn ergibt.¹⁰⁷ Bei dem invertierten Verb handelt es sich in der Regel um

⁹⁸ Rey-Debove & Rey (1993), S. 1147. Siehe auch die digitalisierte Fassung des *Dictionnaire de Trévoux*: <http://archive.org/stream/dictionnaireuniv05fure#page/117/mode/1up>.

⁹⁹ Vgl. Bally (1944), S. 214.

¹⁰⁰ Zitiert nach Grevisse & Goosse (1993), S. 573.

¹⁰¹ Grevisse & Goosse (2008), S. 473.

¹⁰² Vgl. Odile Halmøy (1983), S.233.

¹⁰³ Vgl. Dessaintes, S.73.

¹⁰⁴ Vgl. Odile Halmøy (1983), S.233.

¹⁰⁵ Vgl. Dessaintes (1960), S. 13.

¹⁰⁶ Vgl. Dessaintes (1960), S. 19f.

¹⁰⁷ Vgl. Dessaintes (1960), S. 73.

ein Redeverb bzw. um ein Verb, das eine Redeabsicht ausdrückt, oder auch ein Meinungsverb.¹⁰⁸ Außerdem gilt:

„Psychologiquement, le sujet inversé est un sujet effacé ; il s’est effacé devant un complément plus important que lui et qui commençait la phrase ; en outre, les verbes ainsi employés sont sémantiquement, assez faibles (ce sont de simples «introduceurs» et la déclaration est plus importante que la tournure qui l’introduit).“¹⁰⁹

Aufgrund ihrer Häufigkeit und ihres mangelnden semantischen Gehalts sieht Dessaintes die *incidente* nur als Diskursmorphem mit der Wertigkeit einer Präposition wie *selon*, *d’après* oder *suivant* an, woran auch die Vielfalt der Konstruktionen¹¹⁰, die innerhalb einer *incidente* möglich sind, nichts ändert.

Noch am Ende der 1970er Jahre bemängelt Prince (1978), dass es innerhalb der Erzähltheorie keinen Terminus für redeeinleitende bzw. redebegleitende Konstruktionen gebe¹¹¹, und schlägt dafür *discours attributif* vor. Dieser umfasst die am häufigsten vertretene Gruppe

«[...] sujet-verbe (ou verbe-sujet) renvoyant au locuteur et à son acte de parole [et] constitu[ant] la forme la plus courante du *discours attributif*“¹¹²

sowie alle Konstruktionen, die den Gesprächspartner näher beschreiben und die über die Art und Weise, wie gesprochen wird – dazu zählen der Umgangston, die verwendete Mimik und Gestik, Situationsbeschreibungen etc. – Auskunft geben. Da der *discours attributif* per Definition nur im Zusammenhang mit der direkten Rede in Erscheinung treten kann, ist er unweigerlich mit der sprechenden Person verbunden und trägt wesentlich zu ihrer Charakterisierung bei¹¹³. Obwohl der *discours attributif* für den Leser von großem Nutzen ist, stellt er doch eine gewisse Distanz her, indem der Erzähler mittels einer Redeeinleitung oder Redebegleitung stärker in den Vordergrund tritt.¹¹⁴

Die Redebegleitung kann, unabhängig von der Länge der zitierten Rede, zwei Positionen einnehmen, jedoch muss sie immer im ersten sich daran anschließenden Satz stehen. Bei nur einem Satz ist sie entweder darin eingebettet (Bsp. 37) oder sie steht am Ende (Bsp. 1444)¹¹⁵:

(37) – *C’est vrai, fit Maud pensive, vous faisiez très bon ménage, toutes les deux.*¹¹⁶

¹⁰⁸ Vgl. Dessaintes (1960), S. 73f.

¹⁰⁹ Dessaintes (1960), S. 74.

¹¹⁰ Auf diese wird in Kapitel 2.2.3 näher eingegangen.

¹¹¹ Vgl. Prince (1978), S. 305.

¹¹² Prince (1978), S. 307.

¹¹³ Vgl. Prince (1978), S. 309.

¹¹⁴ Vgl. Prince (1980b), S. 271.

¹¹⁵ Vgl. Grevisse & Goosse (1993), S. 573. bzw. Riegel *et al.* (2009), S. 1011.

¹¹⁶ Marcel Prévost, *Les demi-vierges* (1894).

(1444) – *Comme vous souffrez ! murmura-t-elle.*¹¹⁷

Die direkte Rede kann auch gänzlich unbegleitet sein. Dies ist vor allem dann der Fall, wenn sich ein Gespräch über mehrere Zeilen hinzieht und für den Leser durch den Zusammenhang klar wird, wer gerade spricht. Dies macht eine erneute Redeeinleitung oder -begleitung überflüssig.¹¹⁸

Der Einschub enthält zwingend ein *verbum communicandi* (entweder *dire* oder eine Variante davon) und ein Subjekt, das dem Verb nachgestellt wird. Der eingeschobene Satz kann kein Objekt enthalten, stattdessen bildet der gesamte Aussagesatz, in den er eingebettet ist, das semantische Objekt des im Einschub enthaltenen Redeverbs.¹¹⁹ Das *verbum communicandi* steht an erster Stelle des eingeschobenen Satzes, solange dieser kein Dativobjekt in Form eines Personalpronomens enthält.

(615) – *Allons, Nanon, haut le pied, dit le bonhomme. Prends-moi cela, ce sera pour le dîner, je régale deux Cruchot.*

(667) – *Mais, cria le président, écoutez-moi donc, alors.*¹²⁰

Handelt es sich bei dem Subjekt um ein Personalpronomen oder das Pronomen *on*, befindet sich dieses zwischen dem Hilfsverb und dem *participe passé*:

(1498) – *Savez-vous une singulière nouvelle ? avait-il dit deux ans auparavant, il y a un autre poète en Bourgogne !*¹²¹

Liegt als Subjekt ein *je* vor, wird in der literarischen Sprache das Endungs-*e* des Verbs zu einem *é* bzw. *è*.¹²² Innerhalb des Korpus wurde jedoch kein Beispiel dafür ermittelt.

Eventuelle Erweiterungen stehen hinter der Verknüpfung aus Verb und Subjekt, mit Ausnahme der Dativpronomina, die vorangestellt werden:

(311) – *Messieurs, leur dis-je, il se trouve dehors un camarade qui veut se chauffer à nos dépens.*¹²³

Zwischen Verb und Subjekt können Adjektive bzw. Partizipien, Gerundive, aber auch Adverbien und adverbiale Wendungen stehen. Allerdings darf das Subjekt dann kein Pronomen sein. Wird das Subjekt durch die Personalpronomina *il*, *elle* oder *on* vertreten und endet das Verb auf einen Konsonanten, wird ein „*t analogique*“, das durch Bindestriche von Verb und Pronomen getrennt ist, eingefügt.¹²⁴

¹¹⁷ Marcel Prévost, *L'automne d'une femme* (1893).

¹¹⁸ Vgl. Riegel *et al.* (2009), S. 1011. Siehe auch Cornulier (2004): «[...] quand le contexte suffit, on se passe souvent d'incise.» (S. 106).

¹¹⁹ Vgl. Odile Halmøy (1983), S.233.

¹²⁰ Die Beispiele (615) und (667) stammen aus *Eugénie Grandet* (1843) von Honoré de Balzac.

¹²¹ Honoré de Balzac, *Les paysans* (1850).

¹²² Vgl. Grevisse & Goosse (2008), S. 473.

¹²³ Honoré de Balzac, *Le médecin de campagne* (1833).

¹²⁴ Vgl. Grevisse & Goosse (2008), S. 473.

Die Rolle des Subjekts kann von einem Personalpronomen, einem Eigennamen, einem Nominalsyntaxema¹²⁵, das sich aus einem Substantiv und dessen Erweiterungen zusammensetzt, oder einem Indefinitpronomen übernommen werden. In Verbindung mit *dire* oder einem Synonym dazu weist das Subjekt immer das Merkmal „menschlich“ auf.

2.2.2 *Verba communicandi* innerhalb des *discours citant*

Während die Arbeiten zum *discours citant* als solchem überschaubar sind, erfreut sich die Analyse der *verba communicandi* größerer Beliebtheit, vor allem auch bei Linguisten.¹²⁶

Die Redeverben nehmen eine besondere Stellung ein, da mit dem Redeverb nicht nur „[...] über eine Rede gesprochen oder kommuniziert [wird], sondern auch über die Sprache der Rede“¹²⁷, die Metasprache. Darüber hinaus erzeugt das Redeverb eine Redundanz, indem es dem Hörer bzw. Leser zusätzlich zu der Ankündigung der direkten Rede durch graphische Markierungen (Doppelpunkt, Anführungszeichen) anzeigt, dass gesprochen wird. Körner (1977) kommt zu dem Schluss, dass die Anzahl der in Verbindung mit direkter Rede stehenden *verba communicandi* ab dem Beginn des 19. Jahrhunderts stetig zunimmt und interpretiert dies als „Bewusstwerdung des wichtigsten linguistischen Merkmals der direkten Redeverben, [...] der metalinguistisch redundanten Funktion“¹²⁸. So sieht er als Ursache für die zunehmende Verwendung von *faire* als *verbum communicandi* die „Erkenntnis der Sprachbenutzer, die Kommunikation reproduzieren, dass direkte Rede vor allem *expression* ist“¹²⁹. Dies gehe sogar so weit, dass *faire* von einzelnen Autoren auch dann als redebegleitendes Verb eingesetzt werde, wenn ein syntaktischer Satz vorliege und ein Verweis auf die sprechende Figur nicht unbedingt nötig wäre.¹³⁰ *Faire* sei somit oftmals nur noch ein Indiz für die inhaltliche Leere einer Äußerung.

Sowohl für die Redebegleitung als auch für die Redeeinleitung gilt, dass das verwendete Redeverb entweder transitiv (vom Typ *dire, écrire, penser*¹³¹) oder intransitiv

¹²⁵ Vgl. Odile Halmøy (1983), S. 237.

¹²⁶ Dies bezeugen die zahlreichen linguistisch ausgerichteten Arbeiten von beispielsweise Chatton (1953), Charolles (1976), Martins-Baltar (1976), Cornulier (1978), Marie-Madeleine de Gaulmyn (1986), Annie Delaveau (1988), Annette Sabban (1988), Elżbieta Jamrozik (1992), Monique Monville-Burston (1993), Bäckvall (1999), Fairon (2004), Elżbieta Biardzka (2007), Ewelina Marczak (2007) oder Lamiroy & Charolles (2008).

¹²⁷ Körner (1977), S. 157.

¹²⁸ Körner (1977), S. 163.

¹²⁹ Körner (1977), S. 164.

¹³⁰ Vgl. Körner (1977), S. 165.

¹³¹ Grevisse & Goosse (2008), S. 519.

(*s'écrier*, *se récrier* oder *s'exclamer*¹³²) ist. Ist ein transitives Redeverb enthalten, ist ein direktes Objekt zwingend notwendig, was entweder durch einen *discours indirect lié* realisiert wird oder aber durch einen *discours direct*, der einem „équivalent non syntaxique, parataxique [...] de cet objet“¹³³ entspricht.

Grundsätzlich kann das *verbum communicandi* jede Zeitform im Indikativ annehmen, jedoch dominieren das *présent* und das *passé simple*. Nicht immer lässt sich das verwendete Tempus eindeutig bestimmen: Ausgerechnet das in der *incise* am häufigsten gebrauchte Redeverb *dire* verweigert sich durch die Homographie seiner Präsens- bzw. *participe-passé*-Form *dit* einer eindeutigen Zuordnung. Odile Halmøy bezeichnet dieses Phänomen als „neutralisation temporelle“¹³⁴. Meist steht das *verbum communicandi* in der dritten Person Singular, doch je nach Situation kommen auch die übrigen Formen zum Einsatz.

Die *verba communicandi*, die synonym zu *dire* verwendet werden können, weisen zwei Hauptmerkmale auf: Eines der Seme ist *dire*, ein anderes drückt die Modalität der Äußerung aus. Erfüllt ein Verb diese beiden Eigenschaften, so kann es – z.B. in Form eines Gerundivums oder einer modalen adverbialen Ergänzung – das Redeverb *dire* inhaltlich erweitern.¹³⁵ Das *verbum communicandi* im eingeschobenen Satz ist nahezu aton, da es nur eine Unterbrechung signalisiert bzw. das Subjekt näher bestimmt.¹³⁶

Sei es wegen besserer Beschreibungsmöglichkeit für die Umstände der direkten Rede, sei es aufgrund größerer Variationsmöglichkeit: Die Zahl der Redeverben für das gegenwärtige Französische hat seit dem 19. Jahrhundert stark zugenommen. Jedoch wurden noch 1935 beispielsweise *ricaner*, *rétorquer* und *souigner* von der *Académie française* nicht als Redeverben eingestuft, und auch im Jahr 2000 wurden einige Verben, darunter *hasarder*, *indiquer*, *interjeter*, *maugréer*, von der *Académie* ebenfalls nicht in diesem Zusammenhang erwähnt, obwohl sie von Literaten als auch Journalisten als solche bereits eingesetzt wurden und werden¹³⁷.

Einige Aufmerksamkeit wurde dem Redeverb *faire* zuteil. Sein Gebrauch geht bis ins Altfranzösische zurück, wo *faire* sowohl in der Redeeinleitung als auch in der Redebegeleitung durchaus üblich war:

*Ça [= ici] n'avez vos, FET il, que fere (Chrétien de Troyes, Erec, 172).*¹³⁸

¹³² Vgl. Grevisse & Goosse (2008), S. 520.

¹³³ Grevisse & Goosse (2008), S. 519.

¹³⁴ Odile Halmøy (1983), S. 237.

¹³⁵ Vgl. Odile Halmøy (1983), S.235.

¹³⁶ Vgl. Odile Halmøy (1983), S.235.

¹³⁷ Vgl. Grevisse & Goosse (2008), S. 520.

¹³⁸ Zitiert nach Grevisse & Goosse (2008), S. 520.

Seit dem 19. Jahrhundert wird *faire* verstärkt als Ersatz für *dire* innerhalb einer Redebegleitung eingesetzt.

(813) – *Je ne sais rien, moi !... fit-il en remuant les épaules, mais Claparon a des connaissances, il a travaillé pour des banquiers de haut bord, et il s'est mis à rire en disant*.¹³⁹

(1755) – *Ah fit-elle, avec une intonation de voix singulière, où il y avait de la douleur et de l'ironie. Ah !*¹⁴⁰

(1312) – *Ah ! si ce n'était que cette promesse, fit-il. Et ainsi je connus qu'autre chose liait les deux jeunes hommes, mais sans pouvoir deviner quoi.*¹⁴¹

Laurence Rosier (1999) bezeichnet *faire* als Kondensat, weil es in Gestalt einer Geste (*le faire*) meist die Art zu reden (*le dire*) abbildet und damit gleich zwei Aspekte des Sprechaktes abdeckt¹⁴².

2.2.3 Konstruktionstypen innerhalb des redebegleitenden *discours citant*

Den Kern eines redebegleitenden *discours citant* bildet die Verknüpfung aus *verbum communicandi* und Subjekt, das dem Verb in der Regel nachgestellt ist. Jedes dieser beiden Elemente kann durch weitere ergänzt werden. Die Möglichkeiten zur Ausgestaltung der Redebegleitung sind zwar vielfältig, aufgrund der Inversion von Subjekt und Verb jedoch nicht unendlich.

Eine für diese Arbeit sehr wichtige Grundlage stellt der Beitrag von Odile Halmøy (1983) dar. Halmøy beschreibt und ordnet das «micro-système»¹⁴³ des *discours citant* in der Redebegleitung. Halmøys Einteilung in verschiedene Gruppen ist insofern in die Programmierung der Software eingeflossen, als sie der Ausarbeitung der für das Erkennen von Textmustern notwendigen Regulären Ausdrücke zugrunde gelegt wurde. In Anlehnung an ihre Ausführungen wurden für die Programmierung von DCAnalyse die nachfolgenden Konstruktionstypen innerhalb des *discours citant* unterschieden. Für alle Untergruppen und deren weitere mögliche Varianten¹⁴⁴ wird jeweils ein eigener Regulärer Ausdruck formuliert.¹⁴⁵

¹³⁹ Honoré de Balzac, *Les petits bourgeois* (1850).

¹⁴⁰ Gustave Flaubert, *L'éducation sentimentale* (1869).

¹⁴¹ Henri Alain-Fournier, *Le grand Meaulnes* (1913).

¹⁴² Vgl. Laurence Rosier (1999), S. 211.

¹⁴³ Odile Halmøy (1983), S. 233.

¹⁴⁴ Je nach Wortstellung muss der Reguläre Ausdruck verändert werden. Für «... », *dit-il* lautet der Ausdruck anders als für «... » *dit le jeune homme*..

¹⁴⁵ Einzelne Beispiele dafür sind in Kapitel 3.3.8 im Detail aufgeführt.

Gruppe 1: Verb + Subjekt (ohne Erweiterung)

Die kürzeste Ausprägung eines solchen Einschubs ist der Typ «dit-il». Wohl auch aufgrund seiner Kürze ist der Basistyp der am häufigsten auftretende.

(356) – *C'est fait, voici le lit, et quelques piqûres ne sont pas des coups de sabre, reprit la brave petite fille.*¹⁴⁶

Da dieser „Basistyp“ verschiedene Ausprägungen hat, wurden für die Suche nach T1 folgende Untergruppen definiert:

T1a: Verb + Personalpronomen

(445) – *Que j'aïlle à son secours, s'écria-t-il, ou que je meure!*¹⁴⁷

T1b: Verb + Nomen

(77) – *Que voulez-vous ? demanda le vieillard.*¹⁴⁸

T1c: Verb + Eigename

(1213) – *Ah ! mon petit Fabrice, s'écria la Fausta ; si je savais où te prendre!*¹⁴⁹

Für Typ 1b und 1c muss außerdem berücksichtigt werden, dass das Nomen mit verschiedenen Attributen versehen werden kann, d.h. der Reguläre Ausdruck muss diese Möglichkeit ebenfalls abbilden.

Gruppe 2: Verb + Subjekt + Subjekterweiterung

Zu dieser Gruppe von Einschüben gehören diejenigen eingeschobenen Sätze, deren Subjekt durch ein attributives oder losgelöstes Adjektiv, ein losgelöstes *participe présent* oder *participe passé*, einen Relativsatz oder eine feste Wendung erweitert werden. In letzterem Falle benennt das erste Element einen Körperteil oder ein damit in Verbindung stehendes Objekt (Kleidung, Accessoires etc.), während das zweite Element in Beziehung zum ersten steht (z.B. *les yeux fermés, les mains dans les poches*).¹⁵⁰ Es ergeben sich daher folgende Untergruppen, nach denen der *discours citant* durchsucht wird:

T2a: Verb + Nomen + Adjektiv / *participe passé*

(1729) – *Elles sont vraies ! répliqua Mme Arnoux impatientée. Car, enfin, tu l'as acheté.*¹⁵¹

¹⁴⁶ George Sand, *La mare au diable* (1846).

¹⁴⁷ Bernardin de Saint-Pierre, *Paul et Virginie* (1787).

¹⁴⁸ Honoré de Balzac, *La peau de chagrin* (1831).

¹⁴⁹ Stendhal, *La chartreuse de Parme* (1839).

¹⁵⁰ Vgl. Odile Halmøy (1983), S. 237 ff.

¹⁵¹ Gustave Flaubert, *L'éducation sentimentale* (1869).

Das Adjektiv bzw. das *participe passé* kann seinerseits um Adverbien, adverbiale Wendungen oder Präpositionalsyntagmen erweitert werden.

T2b: Verb + Nomen + Adjektiv / *participe présent*

(2007) – *Voyons ! fit Maud haussant les épaules, tout cela est très ennuyeux, certes ; mais ce n'est pas une raison pour ne pas revoir Paul, qui t'aime, que tu aimes.*¹⁵²

Dem *participe présent* folgt meist eine Erweiterung, auch wenn es keine feste Regel dafür gibt. Ist das Verb des Partizips ein dynamisches¹⁵³, wäre stattdessen auch eine *gérondif*-Konstruktion möglich. In diesem Fall würde es sich jedoch um eine Erweiterung des Verbs und nicht des Subjekts handeln, was wiederum einen, wenn auch nur leichten, Bedeutungsunterschied zur Folge hätte.

T2c: Verb + Nomen + Relativsatz

(1066) – *Oui ! oui ! sans doute, répondit Michel Ardan qui l'examinait en artiste.*¹⁵⁴

(975) – *Alors, messieurs, vous ne vous opposerez pas à ce que nous exécutions les ordres que nous avons reçus ? demanda celui qui paraissait le chef de l'escouade.*¹⁵⁵

(912) – *La justice l'effraie, dit Michu dont les yeux lançaient des flammes comme ceux d'un lion pris dans un filet.*¹⁵⁶

Bei dem Subjekt des Einschubs kann es sich immer nur um ein Nomen handeln.¹⁵⁷ Außer den Relativpronomina *qui* und *que* sind außerdem die Relativa *celui/celle/ceux qui/que* und *dont* möglich.

Gruppe 3: Verb + Subjekt + Verberweiterung

In dieser Gruppe kann das *verbum dicendi* diverse Erweiterungen erfahren. Am häufigsten vertreten sind dabei Adverbien auf *-ment*, *gérondif*-Konstruktionen und Präpositionalsyntagmen. Eine Sonderstellung nimmt dabei die Dativergänzung ein.

T3a: Verb + Nomen + Adverb

(1152) – *Tiens, ma fille, ne nous brouillons pas pour un coffre. Prends donc ! s'écria vivement le tonnelier en jetant la toilette sur le lit.*¹⁵⁸

¹⁵² Marcel Prévost, *Les demi-vierges* (1894).

¹⁵³ Odile Halmøy (1983), S. 239.

¹⁵⁴ Jules Verne, *De la terre à la lune* (1865).

¹⁵⁵ Alexandre Dumas, *Les trois mousquetaires* (1844).

¹⁵⁶ Honoré de Balzac, *Une ténébreuse affaire* (1846).

¹⁵⁷ Vgl. Odile Halmøy (1983), S. 239.

¹⁵⁸ Honoré de Balzac, *Eugénie Grandet* (1843).

Auf *-ment* endende Adverbien stehen entweder zwischen dem Verb und seinem Subjekt oder hinter der Verbindung aus Verb und Subjekt. Die Stellung zwischen Verb und Subjekt ist jedoch nur dann möglich, wenn das Subjekt nicht durch ein Pronomen ausgedrückt wird.¹⁵⁹

T3b: Verb + Nomen + *gérondif*-Konstruktion

(3470) – *Et peut-être aussi, ajouta le banquier en souriant, grâce un peu à mon influence.*¹⁶⁰

T3c: Verb + Nomen + Präpositionalsyntaxma

(3209) – *Croyez-vous que nous irons plus vite avec chacun deux chevaux ? demanda Planchet avec son air narquois.*¹⁶¹

(875) – *Vous ! fit-elle avec étonnement. Mais je vous croyais très gai ?*¹⁶²

Das Präpositionalsyntaxma kann mit unterschiedlichen Präpositionen gebildet werden, deren Häufigkeit jedoch stark variiert. Sehr häufig kommt *avec* zur Verwendung. Für diese Konstruktionen gilt, dass *avec* immer nur mit einem Substantiv kombiniert wird und dieses ohne bestimmten oder unbestimmten Artikel bzw. Possessivpronomen steht, sofern es nicht näher bestimmt wird, aber mit Artikel bzw. Possessivpronomen, sobald das Substantiv eine attributive Ergänzung bekommt.¹⁶³

Da *à* und *de* meistens in Verbindung mit bestimmten Substantiven (*voix, ton, air*) und außerdem recht häufig gebraucht werden, wurden diese einer weiteren Gruppe (Gruppe 4) zugeordnet.

T3e: Verb + Nomen + Dativerweiterung

(2905) – *Qu'est-ce cela, je vous prie, M. le lieutenant ? demanda-t-il à l'officier chargé de l'escorter.*¹⁶⁴

(290) – *Vous l'embrassez trop fort, lui dit Marie en repoussant doucement la tête du laboureur, vous allez le réveiller.*¹⁶⁵

Eine Besonderheit des Dativobjekts besteht in seiner Fähigkeit, selbst den Kern zusätzlicher Erweiterungen zu bilden. Wird das Objekt allerdings pronominal ausgedrückt, z.B. durch *lui* oder *leur*, steht das Pronomen an erster Stelle des Einschubs und verliert dadurch diese Fähigkeit.¹⁶⁶

¹⁵⁹ Vgl. Odile Halmøy (1983), S. 241.

¹⁶⁰ Gustave Flaubert, *L'éducation sentimentale* (1869).

¹⁶¹ Alexandre Dumas, *Les trois mousquetaires* (1844).

¹⁶² Gustave Flaubert, *Madame Bovary* (1857).

¹⁶³ Vgl. Odile Halmøy (1983), S. 242.

¹⁶⁴ Alexandre Dumas, *La tulipe noire* (1850).

¹⁶⁵ George Sand, *La mare au diable* (1846).

¹⁶⁶ Vgl. Odile Halmøy (1983), S. 243.

Gruppe 4: Diverse Konstruktionen

TKonj: Verb + Subjekt + Nebensatz

(999) – *Mais quelle diable de vilénie avez-vous donc faite là ? dit Porthos lorsque l'alguazil en chef eut rejoint ses compagnons, et que les quatre amis se retrouvèrent seuls.*¹⁶⁷

Tvoix: Verb + Subjekt + voix/ton/air

(254) – *Jamais nous n'arriverons à Vierzon pour le train, dit-il à mi-voix.*¹⁶⁸

(1296) – *Madame n'a plus besoin de pain ? demanda-t-elle d'une voix rauque.*¹⁶⁹

Die Präposition *à* findet sich fast nur in Verbindung mit *voix*, dem wiederum eine adjektivische Erweiterung vorausgeht, z.B. mittels *mi-*, oder nachfolgt, z.B. *haute* oder *basse*. Steht *voix* hingegen in Kombination mit anderen Adjektiven, so kommt die Präposition *de* zum Einsatz. *de* findet sich in wenigen festen Wendungen, die in der Regel die Substantive *voix*, *ton* oder *air* enthalten:¹⁷⁰

(1341) – *Et c'est cette liqueur-là qui va faire sauter nos rochers ? dit Pencroff d'un air assez incrédule.*¹⁷¹

(2940) – *Je puis sortir ? demanda-t-elle d'un ton bref.*¹⁷²

Toreille: Verb + Subjekt + oreille

(2736) – *Hein ! qu'est-ce que je vous ai dit ? fit le garde historien à l'oreille du patient.*¹⁷³

¹⁶⁷ Alexandre Dumas, *Les trois mousquetaires* (1844).

¹⁶⁸ Henri Alain-Fournier, *Le grand Meaulnes* (1913).

¹⁶⁹ Émile Zola, *La bête humaine* (1890).

¹⁷⁰ Vgl. Odile Halmøy (1983), S. 242.

¹⁷¹ Jules Verne, *L'île mystérieuse* (1875).

¹⁷² Émile Zola, *Au bonheur des dames* (1883).

¹⁷³ Alexandre Dumas, *La tulipe noire* (1850).

Kapitel 3

Datengrundlage und Methodik

Für die angestrebte teilautomatisierte quantitative Analyse des *discours citant* benötigt man ein ausreichend großes Korpus und ein Programm, das in der Lage ist, die im Korpus enthaltenen Texte so zu verarbeiten, dass eine Datengrundlage entsteht, die ohne weiteren größeren Aufwand ausgewertet werden kann.

3.1 Das Arbeitskorpus

3.1.1 Quellen

Die im Arbeitskorpus enthaltenen französischen E-Texte stammen ausnahmslos aus im Internet frei zugänglichen Datenbanken. Für die vorliegende Arbeit wurden vor allem die umfangreichen Bestände der digitalen Bibliothek *Gallica*¹, ein Projekt der Bibliothèque nationale de France (BnF), der Organisation *Project Gutenberg*² mit Sitz in den Vereinigten Staaten von Amerika sowie der *Association des bibliophiles universels*³ (ABU) konsultiert. Alle drei gestatten den kostenlosen Download sowie die Benutzung der Volltexte und bieten außerdem für viele Texte verschiedene Dateiformate an, was vor allem dann nützlich sein kann, wenn z.B. eine Textdatei Formatierungsfehler aufweist, die in der HTML-Version nicht vorkommen. Frei erhältlich sind nur diejenigen Romane, deren Urheberrecht⁴ abgelaufen ist.

¹ <http://gallica.bnf.fr/>.

² <http://www.gutenberg.org/>.

³ <http://abu.cnam.fr/>.

⁴ Zur Anwendung kommt das Recht des Landes, in dem sich der Server befindet, der die Dateien zum Download anbietet. In Frankreich ist das Urheberrecht 70 Jahre nach dem Tod des Autors abgelaufen. *Project Gutenberg* verfährt aufgrund des für die USA zur Anwendung kommenden Copyrights so, dass "all works whose copyrights were secured before 1923 are now in the public domain, regardless of where they were published" (http://www.gutenberg.org/wiki/Copyright#Rule_1).

3.1.2 Qualität der Quellen

Die Qualität der einzelnen E-Texte – gemeint ist damit die Art und Weise, wie die Texte formatiert sind – variiert zuweilen sehr stark. Die Bandbreite reicht dabei von nahezu perfekter bis hin zu mangelhafter Formatierung. Die Ursache hierfür liegt in dem komplizierten und zudem kostspieligen Verfahren, das ein Buch bei der Digitalisierung durchläuft. Zunächst wird ein Buch Seite für Seite eingescannt, was z.B. in der Bibliothèque nationale de France mithilfe spezieller Hardware weitgehend automatisiert erfolgt. Eine OCR⁵-Software wandelt in einem weiteren Schritt alle Bilddateien in eine Textdatei um. Ob dieser Formatwechsel erfolgreich, d.h. fehlerfrei, vollzogen werden kann, hängt von mehreren Faktoren ab, wie z.B. dem allgemeinen Zustand des Buches (Flecken, Knicke, Papierstärke), aber auch von den verwendeten Schriftarten, Schriftgrößen und Sonderzeichen.

„Pour éliminer complètement ces erreurs, il faudrait avoir recours à un opérateur qui les corrige manuellement, ce qui rendrait la numérisation 10 à 25 fois plus coûteuse : un choix impossible lorsque l'on vise la numérisation de masses importantes de documents. Toutefois, une sélection d'ouvrages accessibles à terme dans Gallica 2 bénéficiera de cette correction.“⁶

Der Kostenfaktor bei der Digitalisierung spielt zumindest bei *Project Gutenberg* eine untergeordnete Rolle, da das Einscannen von Textseiten und die anschließende Texterkennung und Korrektur durch viele freiwillige Helfer erfolgt:

“Project Gutenberg is not powered by financial or political power. Therefore Project Gutenberg is powered totally by volunteers. [...] We offer as many freedoms to our volunteers as possible, in choices of what books to do, what formats to do them in, or any other ideas they may have concerning ‘the creation and distribution of eBooks’.”⁷

Aber gerade die Beteiligung vieler unterschiedlicher Personen an der Digitalisierung eines Buches kann mit ein Grund für die teilweise sehr heterogene Qualität der auf diese Weise gewonnenen E-Texte sein.

Dass sich die Formatierung der Digitalisate weniger an der späteren elektronischen Verarbeitbarkeit orientiert, sondern vielmehr ein Mittelweg gewählt wird, der auf die Bedürfnisse einer breiten Anwenderschicht ausgerichtet ist, zeigt ein genauerer Blick in verschiedene Dateien: In reinen Textdateien kommt es immer wieder vor, dass um der Lesbarkeit im Web-Browser – Mozilla Firefox, Internet Explorer, Safari, Opera u.a. zei-

⁵ Englisch für *Optical Character Recognition*.

⁶ <http://gallica.bnf.fr/html/aide/contenu.html>. Der Link ist inzwischen nicht mehr gültig. Im Internet-Archiv „*Wayback-machine*“ (<http://archive.org/web/web.php>) finden sich leider nur andere, auf derselben Ebene von «*Aide*» liegende Dokumente, z.B. <http://web.archive.org/web/20100218135937/http://gallica.bnf.fr/html/aide/consulter.html>.

⁷ http://www.gutenberg.org/wiki/Gutenberg:Project_Gutenberg_Mission_Statement_by_Michael_Hart.

gen grundsätzlich den Inhalt jeder textbasierten Datei (z.B. *.txt, *.csv und natürlich *.html an – willen ein Zeilenumbruch nach einer bestimmten Zeilenlänge erfolgt. Gleiches lässt sich für HTML-Dateien beobachten, obwohl sich hier eine beliebige Zeilenbreite ganz einfach durch die Verwendung von Stylesheets umgehen ließe und damit der Quelltext von diesem für dieses Dateiformat eigentlich unnötigen Eingriff bewerkstelligen ließe. Grundsätzlich gilt daher: Ein Text, der für den Leser am Bildschirm vermeintlich ansprechend formatiert aussieht, ist nicht unbedingt auch für eine automatisierte elektronische Weiterverarbeitung geeignet und bedarf möglicherweise weiterer, meist manuell vorzunehmender Eingriffe.

3.1.3 Verwendete Werke

Das für die Untersuchung verwendete Korpus umfasst 70 Prosatexte von 16 Autoren, deren Werke zwischen 1747 und 1922 entstanden sind. Insgesamt enthält das Korpus 9,8 Millionen Tokens. Geht man davon aus, dass auf einer durchschnittlichen DIN-A4-Seite ca. 400 Wörter Platz finden, ergibt sich daraus eine Seitenzahl von rund 24550 Seiten.

Bei der Zusammenstellung der E-Texte wurde auf eine möglichst gleichmäßige zeitliche Verteilung der Romane geachtet. Lücken in der Verteilung kommen allerdings dort zustande, wo es für einen Zeitraum zwar passende Werke gäbe, diese jedoch noch nicht digitalisiert wurden und somit nicht als E-Text vorliegen.

Tabelle 3.1 listet alle enthaltenen Romane auf, unter Angabe des Autors, des Erscheinungsjahrs, des Titels, der Quelle der elektronischen Fassung, des Kurztitels, wie er für die Beschriftung von Abbildungen und Tabellen in Kapitel 4 benötigt wird sowie einer Nummer, die ebenfalls bei der Beschriftung von Schaubildern in Kapitel 4 zum Einsatz kommt.

Autor	Jahr	Titel	Quelle	Kurztitel	Nr.
Voltaire	1747	Zadig	BNF	VOL_ZADIG	1
Voltaire	1750	Memnon	BNF	VOL_MEMNON	2
Voltaire	1759	Candide	BNF	VOL_CANDID	3
Voltaire	1764	Le Blanc et le noir	BNF	VOL_BLANCN	4
Voltaire	1767	L'Ingénu	BNF	VOL_INGENU	5
Voltaire	1774	La princesse de Babylone	BNF	VOL_PRINCB	6
Bernardin de Saint-Pierre	1787	Paul et Virginie	BNF	SPI_PAUVIR	7
Victor Hugo	1823	Han d'Islande	Gutenberg	HUG_HANISLA	8
Stendhal	1830	Le rouge et le noir	Gutenberg	STE_ROUGEN	9
Honoré de Balzac	1831	La peau de chagrin	BNF	BAL_PEAUCH	10
...

Tabelle 3.1: Übersicht über die im Korpus enthaltenen Texte.

Autor	Jahr	Titel	Quelle	Kurztitel	Nr.
Victor Hugo	1831	Notre-Dame de Paris	Gutenberg	HUG_NODAME	11
George Sand	1832	Indiana	Bibliopolis	SAN_INDIAN	12
Honoré de Balzac	1833	Le médecin de campagne	BNF	BAL_MEDCAM	13
Honoré de Balzac	1835	La duchesse de Langeais	BNF	BAL_DUCHES	14
Honoré de Balzac	1835	La fille aux yeux d'or	BNF	BAL_FILLEY	15
Honoré de Balzac	1837	César Birotteau	BNF	BAL_CESBIR	16
Stendhal	1839	La chartreuse de Parme	Gutenberg	STE_CHARTR	17
Honoré de Balzac	1842	La femme de trente ans	BNF	BAL_FEMMET	18
George Sand	1842	Consuelo	Gutenberg	SAN_CONSUE	19
Eugène Sue	1842	Les mystères de Paris	Gutenberg	SUE_MYSTER	20
Honoré de Balzac	1843	Eugénie Grandet	BNF	BAL_EUGENI	21
Honoré de Balzac	1843	Le Père Goriot	BNF	BAL_PGORIO	22
Honoré de Balzac	1844	Le colonel Chabert	BNF	BAL_COLCHA	23
Honoré de Balzac	1844	Le lys dans la vallée	BNF	BAL_LYSVAL	24
Alexandre Dumas	1844	Les trois mousquetaires	Gutenberg	DUM_MOUSQU	25
Honoré de Balzac	1845	Le curé de village	BNF	BAL_CUREVI	26
Eugène Sue	1845	Le juif errant	Gutenberg	SUE_JUIFET	27
Honoré de Balzac	1846	Une ténébreuse affaire	BNF	BAL_TENAFF	28
Alexandre Dumas	1846	Le comte de Monte Christo	Gutenberg	DUM_CHRIST	29
George Sand	1846	La mare au diable	Gutenberg	SAN_MAREDI	30
Honoré de Balzac	1847	Le cousin Pons	BNF	BAL_COUSPO	31
Honoré de Balzac	1848	La cousine Bette	BNF	BAL_COUSBE	32
Alexandre Dumas	1849	Le collier de la reine	Gutenberg	DUM_COLLIE	33
Honoré de Balzac	1850	Les paysans	BNF	BAL_PAYSAN	34
Honoré de Balzac	1850	Les petits bourgeois	BNF	BAL_PBOURG	35
Alexandre Dumas	1850	La tulipe noire	Gutenberg	DUM_TULIPE	36
George Sand	1850	François le Champi	ABU	SAN_FCHAMP	37
Gustave Flaubert	1857	Madama Bovary	ABU	FLA_BOVARY	38
Gustave Flaubert	1862	Salammbô	ABU	FLA_SALAMM	39
Victor Hugo	1862	Les misérables	Gutenberg	HUG_MISERA	40
Jules et Edmond de Goncourt	1865	Germinie Lacerteux	Gutenberg	GON_GERMIN	41
Jules Verne	1865	De la terre à la lune	Gutenberg	VER_TERLUN	42
Émile Zola	1867	Thérèse Raquin	Gutenberg	ZOL_THRAQU	43
Gustave Flaubert	1869	L'éducation sentimentale	ABU	FLA_EDUCAT	44
Jules Verne	1869	Vingt-mille lieues sous les mers	Gutenberg	VER_LIEUES	45
Émile Zola	1871	La curée	Gutenberg	ZOL_CUREE	46
Jules Verne	1873	Le tour du monde en quatre-vingt jours	Gutenberg	VER_TOURMO	47
Émile Zola	1873	Le ventre de Paris	Gutenberg	ZOL_VENTRE	48
Émile Zola	1874	La conquête de Plassans	Gutenberg	ZOL_CONQUE	49
Jules Verne	1875	L'île mystérieuse	Gutenberg	VER_ILEMYS	50
Émile Zola	1876	Son Excellence Eugène Rougon	Gutenberg	ZOL_EXCELL	51
Émile Zola	1877	L'assommoir	Gutenberg	ZOL_ASSOMM	52
Émile Zola	1878	Une page d'amour	Gutenberg	ZOL_PAGEAM	53
Émile Zola	1879	Nana	Gutenberg	ZOL_NANA	54
Gustave Flaubert	1881	Bouvard et Pécuchet	ABU	FLA_BOUVAR	55
Émile Zola	1883	Au bonheur des dames	Gutenberg	ZOL_BONHEU	56
...

Tabelle 3.1: Übersicht über die im Korpus enthaltenen Texte.

Autor	Jahr	Titel	Quelle	Kurztitel	Nr.
Émile Zola	1888	Germinal	Gutenberg	ZOL_GERMIN	57
Émile Zola	1887	La terre	Gutenberg	ZOL_TERRE	58
Émile Zola	1890	La bête humaine	Gutenberg	ZOL_BETEHU	59
Émile Zola	1891	L'argent	Gutenberg	ZOL_ARGENT	60
Émile Zola	1892	La débâcle	Gutenberg	ZOL_DEBACL	61
Marcel Prévost	1893	L'automne d'une femme	Gutenberg	PREV_AUTOM	62
Émile Zola	1893	Le docteur Pascal	Gutenberg	ZOL_DRPASC	63
Marcel Prévost	1894	Les demi-vierges	Gutenberg	PREV_VIERG	64
Émile Zola	1894	Les trois villes	Gutenberg	ZOL_TRVIL	65
Henri Alain-Fournier	1913	Le Grand Meaulnes	Gutenberg	AFOU_GD_ME	66
Marcel Proust	1913	Du côté de chez Swann	Gutenberg	PROU_SWANN	67
Marcel Proust	1918	À l'ombre des jeunes filles en fleurs	Gutenberg	PROU_OMBRE	68
Colette	1920	Chéri	Gutenberg	COLE_CHERI	69
Marcel Proust	1922	Le côté de Guermantes	Gutenberg	PROU_GUERM	70

Tabelle 3.1: Übersicht über die im Korpus enthaltenen Texte. Jeder Roman hat einen Kurztitel (zweite Spalte von rechts) sowie eine Nummer (letzte Spalte), die vor allem in den in Kapitel 4 verwendeten Schaubildern zur Identifizierung dienen.

3.2 Für die Untersuchung des *discours citant* verwendete *verba communicandi*

Unter die Bezeichnung *verba communicandi* bzw. Redeverbren fallen alle in der Rede-einleitung und Redebegleitung möglichen Einzelverbren. Es wird auf eine weitere Klassifikation der Redeverbren verzichtet, wie sie andere bereits nach sehr unterschiedlichen Gesichtspunkten vorgenommen haben.⁸ Berücksichtigt werden die nachfolgenden 439 *verba communicandi*:

<i>aboyer</i>	<i>affirmer</i>	<i>approuver</i>	<i>babiller</i>	<i>bramer</i>
<i>accepter</i>	<i>ajouter</i>	<i>appuyer</i>	<i>bafouiller</i>	<i>bredouiller</i>
<i>accorder</i>	<i>alarmer</i>	<i>arguer</i>	<i>balbutier</i>	<i>cacarder</i>
<i>accoucher</i>	<i>alléguer</i>	<i>articuler</i>	<i>barrir</i>	<i>calculer</i>
<i>accuser</i>	<i>amuser</i>	<i>assurer</i>	<i>bavarder</i>	<i>calomnier</i>
<i>achever</i>	<i>annoncer</i>	<i>attaquer</i>	<i>bégayer</i>	<i>cancaner</i>
<i>acquiescer</i>	<i>annuler</i>	<i>attester</i>	<i>bêler</i>	<i>caqueter</i>
<i>adjurer</i>	<i>ânonner</i>	<i>autoriser</i>	<i>bêtifier</i>	<i>caractériser</i>
<i>admettre</i>	<i>apostropher</i>	<i>avancer</i>	<i>beugler</i>	<i>causer</i>
<i>admirer</i>	<i>appeler</i>	<i>avertir</i>	<i>blaguer</i>	<i>certifier</i>
<i>admonester</i>	<i>apprécier</i>	<i>aviser</i>	<i>bougonner</i>	<i>chanter</i>
<i>adresser</i>	<i>apprendre</i>	<i>avouer</i>	<i>brailler</i>	<i>chantonner</i>

⁸ In diesem Zusammenhang sind die Beiträge von Charolles (1976), Martins-Baltar (1976), Bessonnat (1994) und Vivès (1998) zu nennen.

<i>chapitrer</i>	<i>déclamer</i>	<i>engager</i>	<i>hurler</i>	<i>montrer</i>
<i>charger</i>	<i>déclarer</i>	<i>énoncer</i>	<i>implorer</i>	<i>moquer</i>
<i>chasser</i>	<i>décommander</i>	<i>enquérir</i>	<i>improviser</i>	<i>moucharder</i>
<i>chevroter</i>	<i>déconseiller</i>	<i>enseigner</i>	<i>imputer</i>	<i>mugir</i>
<i>chicaner</i>	<i>decréter</i>	<i>entendre</i>	<i>inciter</i>	<i>murmurer</i>
<i>chuchoter</i>	<i>décrier</i>	<i>entêter</i>	<i>incriminer</i>	<i>narrer</i>
<i>chuintier</i>	<i>décrire</i>	<i>épancher</i>	<i>inculper</i>	<i>nasiller</i>
<i>citer</i>	<i>dédire</i>	<i>épiloguer</i>	<i>indigner</i>	<i>négocier</i>
<i>claironner</i>	<i>défendre</i>	<i>époumoner</i>	<i>indiquer</i>	<i>nier</i>
<i>clamer</i>	<i>définir</i>	<i>ergoter</i>	<i>informer</i>	<i>noircir</i>
<i>coasser</i>	<i>dégoiser</i>	<i>éructer</i>	<i>injurier</i>	<i>noter</i>
<i>colporter</i>	<i>délibérer</i>	<i>esclaffer</i>	<i>inquiéter</i>	<i>notifier</i>
<i>commander</i>	<i>délirer</i>	<i>esquisser</i>	<i>insinuer</i>	<i>objecter</i>
<i>commencer</i>	<i>demander</i>	<i>estimer</i>	<i>insister</i>	<i>observer</i>
<i>communiquer</i>	<i>démentir</i>	<i>étonner</i>	<i>instruire</i>	<i>obstiner</i>
<i>compléter</i>	<i>dénigrer</i>	<i>évoquer</i>	<i>insulter</i>	<i>offrir</i>
<i>complimenter</i>	<i>dénoncer</i>	<i>exalter</i>	<i>insurger</i>	<i>opiner</i>
<i>compter</i>	<i>dépeindre</i>	<i>examiner</i>	<i>interdire</i>	<i>opiner</i>
<i>concéder</i>	<i>déplorer</i>	<i>exclamer</i>	<i>interpeller</i>	<i>opposer</i>
<i>conclure</i>	<i>détailler</i>	<i>excuser</i>	<i>interroger</i>	<i>ordonner</i>
<i>condamner</i>	<i>déraisonner</i>	<i>exiger</i>	<i>interrompre</i>	<i>palabrer</i>
<i>conférer</i>	<i>désapprouver</i>	<i>expliquer</i>	<i>intervenir</i>	<i>papoter</i>
<i>confesser</i>	<i>désavouer</i>	<i>exploser</i>	<i>intimer</i>	<i>parlementer</i>
<i>confier</i>	<i>désespérer</i>	<i>exposer</i>	<i>invectiver</i>	<i>parler</i>
<i>confirmer</i>	<i>désoler</i>	<i>exprimer</i>	<i>inviter</i>	<i>peindre</i>
<i>congédir</i>	<i>dessiner</i>	<i>extasier</i>	<i>ironiser</i>	<i>penser</i>
<i>congratuler</i>	<i>détailler</i>	<i>exulter</i>	<i>japper</i>	<i>permettre</i>
<i>conjurier</i>	<i>développer</i>	<i>faire</i>	<i>jaser</i>	<i>pérorer</i>
<i>conseiller</i>	<i>deviser</i>	<i>féliciter</i>	<i>jeter</i>	<i>persifler</i>
<i>consentir</i>	<i>dévoiler</i>	<i>flûter</i>	<i>jubiler</i>	<i>persuader</i>
<i>consoler</i>	<i>dialoguer</i>	<i>formuler</i>	<i>juger</i>	<i>pester</i>
<i>conspuer</i>	<i>dicter</i>	<i>fredonner</i>	<i>jurier</i>	<i>piailler</i>
<i>constater</i>	<i>diffamer</i>	<i>fulminer</i>	<i>lâcher</i>	<i>piauler</i>
<i>conter</i>	<i>dire</i>	<i>garantir</i>	<i>lancer</i>	<i>plaider</i>
<i>contester</i>	<i>disconvenir</i>	<i>gazouiller</i>	<i>larmoyer</i>	<i>plaindre</i>
<i>continuer</i>	<i>discréditer</i>	<i>geindre</i>	<i>lire</i>	<i>plaisanter</i>
<i>contredire</i>	<i>discuter</i>	<i>gémir</i>	<i>louer</i>	<i>plastronner</i>
<i>convaincre</i>	<i>dissenter</i>	<i>glapir</i>	<i>mâchonner</i>	<i>pleurnicher</i>
<i>convenir</i>	<i>dissuader</i>	<i>glisser</i>	<i>mander</i>	<i>polémiquer</i>
<i>converser</i>	<i>distinguer</i>	<i>gloser</i>	<i>marmonner</i>	<i>pontifier</i>
<i>convoquer</i>	<i>divaguer</i>	<i>glouglouter</i>	<i>marmotter</i>	<i>postuler</i>
<i>corner</i>	<i>divulguer</i>	<i>glousser</i>	<i>marteler</i>	<i>pouffer</i>
<i>corriger</i>	<i>éclairer</i>	<i>gribouiller</i>	<i>maugréer</i>	<i>poursuivre</i>
<i>couiner</i>	<i>éclater</i>	<i>grincer</i>	<i>médire</i>	<i>prêcher</i>
<i>couper</i>	<i>écrier</i>	<i>grogner</i>	<i>menacer</i>	<i>préciser</i>
<i>criailler</i>	<i>écrire</i>	<i>grommeler</i>	<i>mender</i>	<i>préconiser</i>
<i>crier</i>	<i>égosiller</i>	<i>gronder</i>	<i>mentionner</i>	<i>prédire</i>
<i>critiquer</i>	<i>émerveiller</i>	<i>gueuler</i>	<i>mentir</i>	<i>prescrire</i>
<i>croquer</i>	<i>émettre</i>	<i>haleter</i>	<i>meugler</i>	<i>présenter</i>
<i>cuisiner</i>	<i>emporter</i>	<i>haranguer</i>	<i>miauler</i>	<i>presser</i>
<i>débiter</i>	<i>empresser</i>	<i>hasarder</i>	<i>minauder</i>	<i>présumer</i>
<i>déblatérer</i>	<i>enchaîner</i>	<i>hennir</i>	<i>moduler</i>	<i>prétendre</i>
<i>décider</i>	<i>enchérir</i>	<i>hoqueter</i>	<i>monologuer</i>	<i>prétexter</i>

<i>prévenir</i>	<i>rappeler</i>	<i>renseigner</i>	<i>ronchonner</i>	<i>stipuler</i>
<i>prier</i>	<i>rapporter</i>	<i>répandre</i>	<i>ronronner</i>	<i>suggérer</i>
<i>proclamer</i>	<i>réciter</i>	<i>repartir</i>	<i>roucouler</i>	<i>supplier</i>
<i>proférer</i>	<i>réclamer</i>	<i>répéter</i>	<i>rouspéter</i>	<i>supposer</i>
<i>professer</i>	<i>recommander</i>	<i>répliquer</i>	<i>rudoyer</i>	<i>susurrer</i>
<i>promettre</i>	<i>reconnaître</i>	<i>répondre</i>	<i>rugir</i>	<i>tancer</i>
<i>prononcer</i>	<i>récrier</i>	<i>repousser</i>	<i>ruminer</i>	<i>taquiner</i>
<i>pronostiquer</i>	<i>récriminer</i>	<i>reprandre</i>	<i>sangloter</i>	<i>targuer</i>
<i>prophétiser</i>	<i>rectifier</i>	<i>représenter</i>	<i>scander</i>	<i>télégraphier</i>
<i>proposer</i>	<i>redire</i>	<i>réprimander</i>	<i>seriner</i>	<i>téléphoner</i>
<i>proscrire</i>	<i>refuser</i>	<i>reprocher</i>	<i>sermonner</i>	<i>tempêter</i>
<i>protester</i>	<i>regimber</i>	<i>réprouver</i>	<i>siffler</i>	<i>tergiverser</i>
<i>prouver</i>	<i>regretter</i>	<i>résoudre</i>	<i>siffloter</i>	<i>tonitruer</i>
<i>psalmodier</i>	<i>réitérer</i>	<i>ressasser</i>	<i>signaler</i>	<i>tonner</i>
<i>quémander</i>	<i>rejeter</i>	<i>résumer</i>	<i>signifier</i>	<i>traduire</i>
<i>questionner</i>	<i>réjouir</i>	<i>rétorquer</i>	<i>solliciter</i>	<i>traiter</i>
<i>rabâcher</i>	<i>relater</i>	<i>rétracter</i>	<i>sommer</i>	<i>trancher</i>
<i>rabrouer</i>	<i>relever</i>	<i>révasser</i>	<i>songer</i>	<i>trionpher</i>
<i>raconter</i>	<i>remarquer</i>	<i>révéler</i>	<i>souffler</i>	<i>trompeter</i>
<i>radoter</i>	<i>remercier</i>	<i>revendiquer</i>	<i>souhaiter</i>	<i>ululer</i>
<i>rager</i>	<i>renâcler</i>	<i>revenir</i>	<i>souligner</i>	<i>vanter</i>
<i>railler</i>	<i>renchérir</i>	<i>ricaner</i>	<i>soupirer</i>	<i>vociférer</i>
<i>raisonner</i>	<i>renier</i>	<i>riposter</i>	<i>spécifier</i>	

Obige Zusammenstellung beruht im Wesentlichen auf der Verbsammlung von Dessaintes (1960) und wurde durch weitere Verben aus den Arbeiten von Dagmar Almenberg (1958), Körner (1977), Elisabeth Gülich (1978), Fónagy (1984), Bessonnat (1994) und Vivienne Mylne (1994) ergänzt.

3.3 Konzeption und Arbeitsweise des Programms DCAnalyse

Nachfolgend werden die wichtigsten Programmabschnitte zusammengefasst, um zu erläutern, wie die für die Auswertung in Kapitel 4 benutzten Daten gewonnen werden. Das Programm besteht aus rund 5000 Zeilen Code der Programmiersprache Perl⁹ und dem Zusatzmodul Perl/TK¹⁰. Für das Tagging der einzelnen Tokens mit Wortarteninformationen wurde auf die Software TreeTagger¹¹ zurückgegriffen. Die Funktionsweise des

⁹ <http://www.perl.org/>.

¹⁰ <http://www.perltk.org/>.

¹¹ <http://www.ims.uni-stuttgart.de/projekte/complex/TreeTagger/>. Siehe dort auch die Installationsanleitung für das Programm.

Programms wird stark vereinfacht in Gestalt eines Flussdiagramms auf Seite 52 in Abbildung 3.2 dargestellt.

3.3.1 Vor dem Start von DCAnalyse

Grundsätzlich gilt: Die Qualität der vom Programm erzeugten Ergebnisse steht und fällt mit der Sorgfalt, die auf die Digitalisierung des zugrundeliegenden gedruckten Textes verwendet wurde. DCAnalyse arbeitet zuverlässig, solange die E-Texte gewisse Mindestanforderungen an die Formatierung erfüllen.

Jede Input-Datei muss nach dem Zeichensatz ISO-8859-1¹², der fast alle Sonderzeichen innerhalb des Französischen abdeckt, kodiert sein. Idealerweise befindet sich innerhalb der Datei jede Äußerung in einem eigenen Absatz, d.h. zu Beginn eines Absatzes steht entweder der redeeinleitende Satz bzw. die direkte Rede selbst, nach deren Ende ein Zeilenumbruch erfolgt. Dieser Idealfall liegt jedoch selten vor, weshalb das Programm in der Lage sein muss, einfache Korrekturen der Formatierung eigenständig vorzunehmen. Das geschieht z.B. dann, wenn sich mehrere Äußerungen in einem Abschnitt befinden oder eine Äußerung nicht am Beginn der Zeile steht. Bei Unregelmäßigkeiten in der Formatierung, die darüber hinausgehen, muss der Anwender jedoch vor Benutzung des Programms den Text von Hand editieren. Die E-Texte, die dieser Arbeit beiliegen, wurden auf diese Schwierigkeiten hin geprüft und gegebenenfalls verbessert, weshalb sie ohne weitere Manipulationen verwendet werden können. Für die Verwendung von Texten anderen Ursprungs empfiehlt es sich jedoch, diese vor der Bearbeitung mit DCAnalyse in einem Texteditor zu öffnen und Formatierungsmängel von Hand zu korrigieren, um später bestmögliche Ergebnisse zu erhalten.

Für den Einsatz von DCAnalyse müssen eine aktuelle Version von Perl, das Modul Perl/TK sowie der Part-of-Speech-Tagger TreeTagger installiert sein. TreeTagger hilft bei der Lemmatisierung und dem Wortarten-Tagging der Tokens. DCAnalyse wurde unter Linux mit der Perl-Version 5.8.8 sowie der Perl/TK-Version 804.027 entwickelt. Alle drei Programme sind für Linux, Windows und MacOS gleichermaßen kostenlos erhältlich, so dass ein plattformübergreifender Einsatz gewährleistet ist.

Für die Dateinamensvergabe der E-Texte gelten folgende Regeln: Ein Dateiname darf als Zeichen Klein- und Großbuchstaben, Unterstriche sowie beliebig viele Ziffern¹³

¹² Dieser Zeichensatz wurde für die Weiterverarbeitung der Input-Dateien gewählt, da diese in der Regel bereits damit formatiert sind und es einen erheblichen Mehraufwand bedeutet hätte, alle Dateien erst noch in den mittlerweile üblich gewordenen Unicode-Zeichensatz UTF-8 zu überführen.

¹³ Regulärer Ausdruck dafür: `[a-zA-Z0-9_]*`.

enthalten, denen ein Punkt sowie die Dateierdung `txt` bzw. `htm` oder `html` folgt. Die Verwendung von Akzentzeichen ist nicht zulässig. Da die Ergebnisse in mehrere automatisch generierte Text- und CSV-Dateien geschrieben werden und dabei der Dateiname der Input-Datei weiterverwendet wird, empfiehlt es sich, diese so zu benennen, dass sowohl Autor als auch der Titel aus dem Dateinamen hervorgehen.

Nach der Installation aller benötigten Programme sind vor dem Start von DCAnalyse noch zwei manuelle Eingriffe seitens des Benutzers nötig. Zum einen müssen in der Datei `dcanalyse.pl` in Zeile 270 und 271 die korrekten Pfadangaben zum Installationsort von TreeTagger angegeben werden. Dieser variiert je nach verwendetem Betriebssystem. Zweitens müssen die zu bearbeitenden Dateien mit einer Anfangs- und einer Endmarkierung versehen werden, die festlegt, welcher Textbereich analysiert wird. Dies erfolgt durch Öffnen der entsprechenden Datei mit einem einfachen Texteditor und Einfügen einer Zeile, bestehend aus dem einzelnen Wort `alpha` am Beginn sowie einer Zeile `omega` am Ende des Prosatextes bzw. des Bereichs, der untersucht werden soll. Der komplette Text zwischen diesen beiden Markern wird abschnittsweise eingelesen, zwischengespeichert und weiterverarbeitet. Durch das Setzen der Markierung wird verhindert, dass die oft vor Beginn des eigentlichen Textes stehenden Informationen zur Benutzung und Weitergabe der Datei sowie die bei HTML-Dateien üblichen Metadaten im Dokumenten-Header in die Analyse einfließen und die Ergebnisse verzerren.

3.3.2 Programmstart

Zum Starten des Programms wechselt man in der Kommandozeile in den Programmordner von DCAnalyse und ruft das Programm mit dem Befehl `dcanalyse.pl` auf. Es erscheint das in Abbildung 3.1 zu sehende Fenster auf dem Bildschirm.

Die Oberfläche bietet mehrere Möglichkeiten zur Dateiauswahl: Der Benutzer kann bequem über einen graphischen Dialog die Texte für die Analyse zusammenklicken. Über den Knopf `Datei(en) wählen` lassen sich eine oder mehrere Dateien auswählen. Über den Knopf `Verzeichnis wählen` werden alle in einem Verzeichnis liegenden Dateien eingelesen. Für die Auswahl können beide Methoden beliebig miteinander kombiniert und gleichzeitig sowohl Text- als auch HTML-Dateien gewählt werden. Das Programm entscheidet aufgrund der Dateierdung, ob Änderungen an der Formatierung vorgenommen werden müssen.

Nach dem Bestätigen des Auswahldialogs werden die gewählten Dateien in einem Teilfenster der GUI angezeigt. Im Falle einer Falschwahl kann die Liste gelöscht und

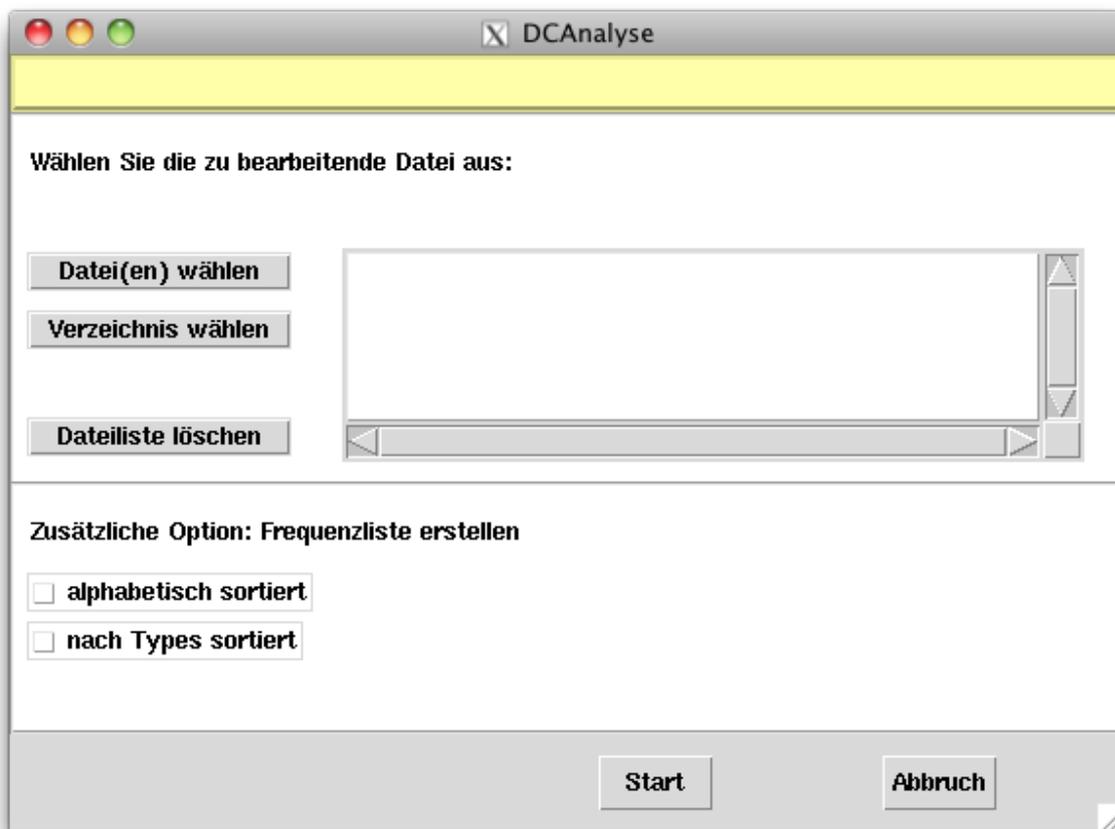


Abbildung 3.1: Programmoberfläche nach Auswahl der zu bearbeitenden Dateien

neu erstellt werden. Zusätzlich besteht die Möglichkeit, sich eine nach Anzahl der Types (über die Checkbox nach Types sortiert) oder eine alphabetisch sortierte Tokenliste (über die Checkbox alphabetisch sortiert) erstellen lassen.

Nach Anklicken des Start-Knopfes bleibt das Programmfenster stehen, und im Hintergrund beginnt die eigentliche, für den Benutzer unsichtbare Verarbeitung des Perl-Skripts. Wurden mehrere Dateien ausgewählt, werden diese nacheinander eingelesen und verarbeitet. Der zwischen dem Anfangs- und Endmarker, alpha bzw. omega, abschnittsweise eingelesene Text wird so zwischengespeichert, dass jeder Absatz als ein Element in einem Array, einer Art Liste, namens @dateiinhalte abgelegt wird. Je nach Eingabeformat, *.txt oder *.htm(1), wird der Eingabetext für die Weiterverarbeitung optimiert, d.h. unnötige Absatzmarkierungen, die zur besseren Lesbarkeit am Bildschirm eingefügt wurden¹⁴, und nicht korrekt verwendete HTML-Tags, die den Text an ungeeigneter Stelle umbrechen, werden entfernt. Anschließend werden leere Elemente, die durch Leerzeilen im Text entstehen, sowie alle weiteren HTML-Tags gelöscht. Die für

¹⁴ Siehe hierzu auch Kapitel 3.4, „Hürden beim Programmieren“.

die korrekte Darstellung von Sonderzeichen benötigten HTML-Entities¹⁵ werden in ISO-8859-1-konforme Zeichen übersetzt¹⁶. Steht die direkte Rede nicht zu Beginn eines Elements, sondern eingebettet oder am Ende, wird der Turn in eine separate Zeile gesetzt und danach als neues Element an dieser Stelle des Arrays eingefügt. Dann wird mit Hilfe einer Kopie dieses Arrays, `@frequenzliste`, jedes Element in seine Einzel-Tokens zerlegt, die Tokens gezählt und, sofern zuvor ausgewählt, Frequenzlisten erstellt. Ob der zu untersuchende Text korrekt eingelesen wurde, kann anhand der Datei `datei_dateiname.txt` überprüft werden. Sind darin Fehler in der Textdarstellung enthalten, sollte die Datei an den entsprechenden Positionen von Hand umformatiert werden, da eine fehlerfreie Verarbeitung von DCAnalyse sonst nicht gewährleistet ist.

3.3.3 Ermitteln der direkten Rede

Nachdem der Dateiinhalt für die nachfolgenden Recherchen eingelesen und aufbereitet wurde, sucht DCAnalyse im Array `@dateiinhalt` nach Elementen, die direkte Rede enthalten. Während dieses Vorgangs werden alle Elemente fortlaufend nummeriert, erkennbar an der zu Anfang jeder Zeile in runden Klammern stehenden Ziffer. Anhand die Nummerierung kann eine einzelne Äußerung leichter ausfindig gemacht werden, und es lässt sich daran schnell ablesen, wie nahe einzelne Dialogstellen beieinanderstehen.

Da beim Einlesen der Datei bereits alle Äußerungen in je ein separates Array-Element überführt wurden, sucht DCAnalyse als nächstes nach einem graphischen Marker für direkte Rede. Als Marker werden in den meisten Fällen französische bzw. englische Anführungszeichen oder ein einfacher bzw. doppelte Bindestriche [" - -- «] verwendet. Hier und da gibt es jedoch auch Abweichungen davon. In diesem Fall sollte der Benutzer vor Benutzung von DCAnalyse den Text umformatieren und den verwendeten Dialogmarker durch einen der vier zuvor genannten ersetzen, da das Programm sonst keine direkte Rede identifizieren kann.

Bei Vorhandensein eines der oben genannten Marker wird zusätzlich überprüft, ob der vorangehende Abschnitt mit einem Doppelpunkt, Komma oder Semikolon endet. Trifft dies zu, wird der letzte ganze Satz dieses Abschnitts zusammen mit der direkten Rede in einem Array namens `@dir_rede_eingeleitet` gespeichert, das pro Element

¹⁵ Innerhalb der *Hypertext Markup Language* (HTML) gibt es einige reservierte Zeichen, an deren Stelle eine Alternativ-Eingabe erfolgen muss. Eine eckige Klammer (<) kann im Quelltext nicht als solche eingegeben werden, sondern muss als `<` übergeben werden, weil der Browser diese sonst als Tag-Markierung identifizieren würde. Innerhalb von HTML können Textelementen mit Hilfe von Tags bestimmte Eigenschaften zugewiesen werden. Aus `<i>kursiv</i>` wird das schräggestellte Wort *kursiv*.

¹⁶ So wird z.B. aus `ç` ein ç oder aus `é` ein é.

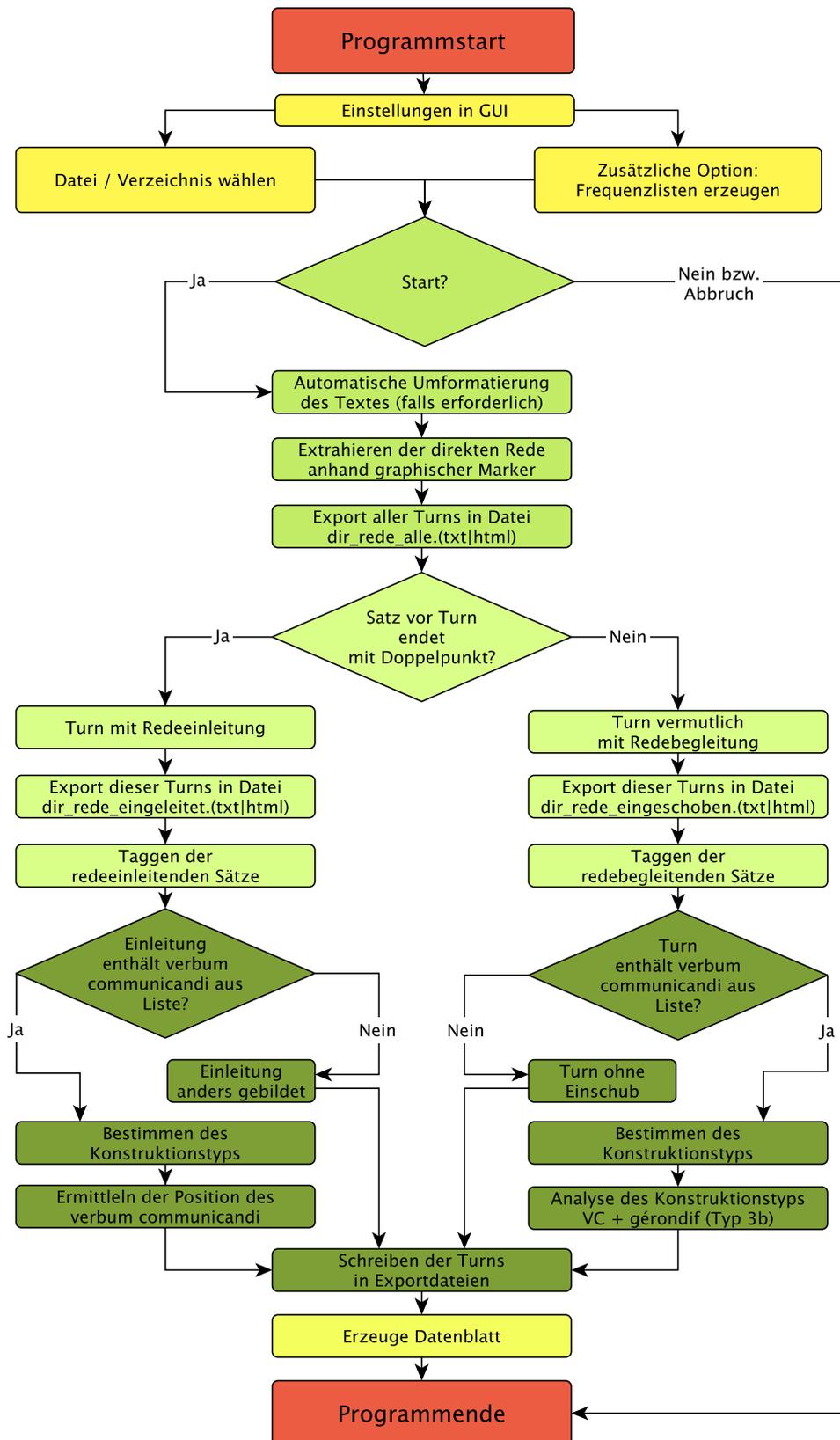


Abbildung 3.2: Schematische Darstellung des Programmablaufs von Anfang bis Ende.

den redeeinleitenden Satz und die direkte Rede bis zum Ende des aktuellen Abschnitts enthält. Ist kein solches Satzzeichen vorhanden, wird der Absatz in ein weiteres Array (@dir_rede_satzanfang) kopiert, das später als Grundlage für die Analyse der redebegleitenden Sätze dienen wird. Gefunden und gesammelt werden Konstruktionen wie die folgenden:

Beispiel für einen Turn mit redeeinleitendem *discours citant*:

(682) *Il dit à Pangloss et à Martin : (683) – Ce bon vieillard me paraît s'être fait un sort bien préférable à celui des six rois avec qui nous avons eu l'honneur de souper.*¹⁷

Beispiel für einen Turn mit redebegleitenden *discours citant*:

(88) – *Eh ! par ici, ma petite ! cria la grosse voix de madame Boche.*¹⁸

Im Verzeichnis /output/dateiname/ werden die Äußerungen später in drei verschiedenen Dateien gespeichert: dir_rede_alle_dateiname.txt enthält alle vom Programm ermittelten Turns, in der Datei dir_rede_eingeleitet_dateiname.txt werden die redeeinleitenden Sätze gespeichert und in dir_rede_satzanfang_dateiname.txt werden die redebegleitenden Sätze abgelegt. Da an dieser Stelle des Programms noch nicht weiter zwischen begleiteten und alleinstehenden Turns differenziert wird, sind in der ersten und letzten dieser drei Dateien auch noch diejenigen Äußerungen enthalten, in denen gar kein *discours citant* vorkommt, wie z.B.

(263) – *Taanach, prends ton nebal et joue tout bas sur la corde d'argent, car mon coeur est triste !*¹⁹

Der alleinstehende *discours citant* wird erst zu einem späteren Zeitpunkt beim Identifizieren der *verba communicandi* vom redebegleitenden *discours citant* getrennt (siehe Kapitel 3.3.7).

3.3.4 Exkurs: Generieren von Listen mit konjugierten Verbformen

Um zu ermitteln, ob es sich bei dem in einem redeeinleitenden bzw. redebegleitenden Satz vorhandenen Verb um ein *verbum communicandi* handelt, benötigt man zum Abgleichen eine Liste mit allen konjugierten Formen der in Frage kommenden Verben. Diese Liste von Hand zu erzeugen, wäre mit großem Arbeitsaufwand verbunden. Zeitsparender und außerdem flexibler ist der Einsatz eines zusätzlichen Perl-Skriptes, da bei

¹⁷ Voltaire: *Candide* (1759). Die Bedeutung der jedem Beispielsatz vorangehenden Ziffern in Klammern wurde bereits in Fußnote 87 in Kapitel 2.2 erläutert sowie zu Beginn dieses Kapitels.

¹⁸ Émile Zola: *L'assommoir* (1877).

¹⁹ Gustave Flaubert: *Salammbô* (1862).

Bedarf einfach neue Verben hinzugefügt werden und die Liste in Sekundenschnelle neu generiert werden kann. Das Zusatzskript erstellt eine Liste mit allen konjugierten Verbformen der in Kapitel 3.2 aufgeführten Verben. Für die Programmierung werden die Verben nach den von Bescherelle (1990) vorgeschlagenen Gruppen zusammengefasst, woraus sich für *dire* und seine Komposita der in Abbildung 3.3 dargestellte Programmcode ergibt.

In dem Array `@vd_typ78` werden die Präfixe der Komposita gesammelt. Für *dire* selbst wird ein leeres Präfix, als "" abgebildet, definiert. Unter Verwendung eines assoziativen Hashes²⁰ wird nun jeweils jedes Element aus dem Array `@vd_typ78` mit jedem Listenelement aller Zeitstufen verknüpft, so dass sich daraus eine vollständige Konjugationsliste für die Verben *contredire*, *dédire*, *dire*, *interdire*, *médire*, *prédire* und *redire* ergibt, in der auch die Infinitive enthalten sind. Der Inhalt dieser Liste wird in verschiedene Textdateien exportiert, z.B. enthält eine Datei nur die konjugierten Formen ohne *participle passé* (`vd_konjug_o_pp.txt`), eine zweite nur die *participle passé*-Formen (`vd_pp.txt`), eine dritte nur die Infinitive (`vd_inf.txt`) und eine vierte nur die Formen des *participle présent* (`vd_ppres.txt`). Diese Verblisten werden von DC Analyse für das Identifizieren von *verba communicandi* in der Redeeinleitung und Redebegleitung benötigt (s. Kapitel 3.3.7). Eine fünfte Datei (`vd_inf_lookup.txt`) enthält pro Zeile nicht nur die konjugierte Verbform, sondern auch den dazugehörigen Infinitiv. Auf diese Weise kann von der konjugierten Form einfach auf den Infinitiv geschlossen werden, was vor allem dann nötig ist, wenn man feststellen möchte, auf wieviele Verb-Types sich die Verb-Tokens verteilen.

3.3.5 Tagging mit der Software TreeTagger

Zwar kann man anhand graphischer Marker die direkte Rede in einer Textdatei recht zuverlässig identifizieren, doch bei der anschließenden Analyse des redeeinleitenden und redebegleitenden *discours citant* stößt man selbst bei der Verwendung Regulärer Ausdrücke schnell an Grenzen, wenn man sich bei der Suche auf Wortformen bzw. Muster beschränkt. Soll beispielsweise bestimmt werden, wie häufig das VC von einem Adverb begleitet wird, so kann man zwar nach der Kombination aus VC und Tokens mit dem Suffix *-ment* suchen, aber man erhält damit nicht diejenigen Adverbien, die nicht nach diesem Schema gebildet werden. Das manuelle Erstellen einer Liste mit allen unregel-

²⁰ Im Gegensatz zu einem Array, dessen einzelne Elemente über ganze Zahlen – das erste Element hat die Indexnummer 0, das zweite die Indexnummer 1 usw. – indiziert werden, wird ein Hash über sogenannte Schlüssel (*keys*) angesprochen. In Abbildung 3.3 werden für den Hash `%zeiten_typ78` Schlüssel erzeugt, die die verschiedenen Tempora abbilden. Jedem dieser Schlüssel wird wiederum eine Liste (Array) von Werten (*values*) mit den entsprechenden Verbformen zugeordnet.

```

#Liste mit VERBA DICENDI vom Typ 78, dire und Komposita (Bescherelle)
@vd_typ78 = ("contre", "dé", "", "inter", "mé", "pré", "re");
%zeiten_typ78 = (
    present           => ["dis","dit","disons","dites","disent"],
    imparfait        => ["disais","disait","disions","disiez","disaient"],
    passesimple      => ["dis","dit","dîmes","dîtes","dirent"],
    futursimple      => ["dirai","diras","dira","dirons","direz","diront"],
    conditionnel     => ["dirais","dirais","dirait","dirions","diriez","diraient"],
    subjonctif       => ["dise","dises","disions","disiez","disent"],
    participepasse   => ["dit","dite","dits","dites"],
    participepresent => ["disant"]);
$inf_typ78 = "dire";
$z=0;
while ($z<@vd_typ78) {
    foreach my $key (keys(%zeiten_typ78)) {
        if($key eq "participepresent") {
            foreach (@{$zeiten_typ78{$key}}) {
                my $form = $vd_typ78[$z].$_;
                push (@vd_ppres, $form);
                $type{$form} = $vd_typ78[$z].$inf_typ78;
            }
        }
        elsif($key ne "participepasse") {
            foreach (@{$zeiten_typ78{$key}}) {
                my $form = $vd_typ78[$z].$_;
                push (@vd_konj_o_pp, $form);
                $type{$form} = $vd_typ78[$z].$inf_typ78;
            }
        }
        else {
            foreach (@{$zeiten_typ78{$key}}) {
                my $form = $vd_typ78[$z].$_;
                push (@vd_pp, $form);
                $type{$form} = $vd_typ78[$z].$inf_typ78;
            }
        }
    }
    my $inf = $vd_typ78[$z].$inf_typ78;
    push(@vd_inf,$inf);
    $z++;
}

```

Abbildung 3.3: Programmcode für das Generieren von konjugierten Verbformen am Beispiel von *dire* und den Komposita *contredire*, *dédire*, *interdire*, *médire*, *prédire* und *redire*. Zunächst wird ein Hash mit allen Verbformen der in Frage kommenden Zeitstufen von *dire* gebildet, indem jedes Element aus dem Array @vd_typ78 mit jedem Listenelement aller Zeitstufen verknüpft wird. Anschließend werden für alle Komposita Listen erstellt, die z.B. nur die Formen des *participe présent* (@vd_ppres), alle konjugierten Formen außer *participe passé* (@vd_konj_o_pp) oder nur Formen mit *participe passé* (@vd_pp) enthalten. Zusätzlich wird eine Liste mit allen Infinitivformen (@vd_inf) erzeugt. Mithilfe dieser Listen kann DCAnalyse an entsprechender Stelle im Programm überprüfen, ob es sich bei einer Verbform um eines der definierten *verba communicandi* handelt.

mäßigen Adverb-Tokens wäre eine Lösung, die jedoch letzten Endes viel zu kompliziert und zeitaufwendig ist und für viele weitere Konstruktionen wiederholt werden müsste.

Eine viel effektivere Methode besteht darin, vorab zu jedem Token des zu untersuchenden Textes die entsprechende Wortart hinzuzufügen, d.h. den Text für die Analyse zu „taggen“, also mit Meta-Informationen zu versehen. Dies ermöglicht neben der Suche nach konkreten Wortformen und Mustern das Auffinden von Abfolgen bestimmter Wortarten, wie z.B. die Abfolge Verb - Personalpronomen im Falle von *dit-il*. Glücklicherweise gibt es bereits Programme, die elektronische Texte automatisch mit morphologischen und syntaktischen Zusatzinformationen, sog. Tags, versehen können. Der Part-of-Speech-Tagger *TreeTagger*²¹ bietet sich für den Einsatz in DC-Analyse an, weil er für alle Betriebssysteme erhältlich und leicht aus einem Perl-Skript heraus auszuführen ist.²² Für den *TreeTagger* sind sog. Tagsets für eine Vielzahl von Sprachen erhältlich, darunter auch eines für Französisch²³, das von Achim Stein vom Institut für Linguistik/Romanistik der Universität Stuttgart zur Verfügung gestellt wird. Die Genauigkeit von *TreeTagger* in Verbindung mit dem französischen Tagset ist mit 95,55 Prozent angegeben²⁴, d.h. dem Programm unterlaufen zwar vereinzelt Fehler beim Bestimmen der Wortarten (siehe Kapitel 3.4.2), was angesichts der dadurch immens verbesserten Suchmöglichkeiten jedoch zu vernachlässigen ist.

Ein grundsätzliches Problem stellt der dem Tagging vorausgehende Schritt der automatischen Tokenisierung dar, bei der ein Text in seine Einzelwörter zerlegt wird: Mehrere Tokens, die eigentlich wie eine Einheit behandelt werden müssten, werden bei der Tokenisierung auseinandergerissen und unabhängig voneinander getaggt, während andere Tokens, für das Französische z.B. die Kontraktion *au* aus *à + le*, getrennt voneinander betrachtet werden müssten. Die Problematik beschreibt Anke Lüdeling (2007) wie folgt:

„In einer sequentiellen Architektur bauen alle weiteren Verarbeitungsschritte auf [der] Tokenisierung auf, d.h. alle Probleme werden ‚mitgeschleppt‘. Die Wortartzuweisung und die Lemmatisierung müssen irgendwie mit merkwürdigen Tokens umgehen. Die Strategien sind hier unterschiedlich. Ganz generell ist es aber so, dass bekannte Probleme [...] einheitlich behandelt werden können, während seltenere Fälle nicht abgefangen werden können [...].“²⁵

Für die vorliegende Arbeit sei hier auf Kapitel 3.4.2 verwiesen, das vereinzelt auftretende

²¹ Download des Programms und weitere Informationen unter <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>.

²² Die Software wurde von Helmut Schmid am Institut für Computerlinguistik der Universität Stuttgart im Rahmen des von 1994-1996 laufenden Projekts „Textcorpora und Erschliessungswerkzeuge“ entwickelt.

²³ Quelle: <http://www.ims.uni-stuttgart.de/~schmid/french-tagset.html>. Die Inhalte dieser Seite wurden übernommen für Tabelle 3.2.

²⁴ <ftp://ftp.ims.uni-stuttgart.de/pub/corpora/french-par-linux.info>.

²⁵ Anke Lüdeling (2007), S. 31.

Fehler beim Tagging erläutert.

Die Programmlaufzeit verlängert sich durch den Einsatz von TreeTagger geringfügig aufgrund der Tokenisierung und des Taggings, was sich allerdings nur bei der Stapelverarbeitung vieler Dateien durch geringfügig längere Wartezeiten bemerkbar macht. Auf einem AMD Athlon 64 X2 Dual Core Prozessor 5200+ betrug die Zeit für das Tagging aller Dialogstellen für eine Datei mit 67118 Tokens weniger als drei Sekunden, für eine Datei mit 275024 Tokens weniger als fünf Sekunden.

Wie ein mit dem TreeTagger ausgezeichnete Satz aussieht, zeigen die folgenden zwei Beispiele:

(490) [, (PUN)][ajouta (VER:simp)][Mlle (NOM)][de (PRP)][Cardoville (NAM)][en (PRP)][souriant (VER:ppre)][avec (PRP)][grâce (NOM)][, (PUN)]²⁶

(3567) [Il (NAM)][regarda (VER:simp)][Mlle (NOM)][de (PRP)][Cardoville (NAM)][d' (PRP)][un (DET:ART)][air (NOM)][presque (ADV)][embarrassé (VER:pper)][et (KON)][RÉPONDIT (VER:simp)][: (PUN)]²⁷

Die Tags lassen sich anhand der Tabelle 3.2 entschlüsseln. Mithilfe der Tags können die Suchmuster der Regulären Ausdrücke so angepasst werden, dass beliebige Abfolgen von Wortarten ermittelt werden können.

3.3.6 Redeeinleitung: Tagging und Ermitteln der *verba communicandi*

Nachdem die Figurenrede, beschrieben in Kapitel 3.3.3, ausfindig gemacht und für die Weiterverarbeitung kopiert wurde, werden als nächstes die redeeinleitenden Sätze von der direkten Rede getrennt. Ein Einleitungssatz kann aus mehreren Teilsätzen bestehen, wie das folgende Beispiel zeigt:

(60) *Elle examina d'un œil indifférent les campagnes du Cher, la Loire et ses îles, Tours et les longs rochers de Vouvray ; puis, sans vouloir regarder la ravissante vallée de la Cise, elle se rejeta promptement dans le fond de la calèche, et dit d'une voix qui en plein air paraissait d'une extrême faiblesse .*²⁸

Ein Blick auf die in den 70 Texten ermittelten redeeinleitenden Sätze zeigt, dass das Redeverb meist sehr nahe bei der direkten Rede steht. Enthält der Einleitungssatz durch

²⁶ Honoré de Balzac: *Eugénie Grandet* (1843).

²⁷ Eugène Sue: *Le juif errant* (1845).

²⁸ Honoré de Balzac: *La femme de trente ans* (1842).

Semikolon voneinander getrennte Teilsätze, kann man sich bei der Suche nach einem *verbum communicandi* auf den letzten Teilsatz vor der Figurenrede beschränken. Von obigem Beispielsatz bleibt dann noch übrig:

(60) ; puis, sans vouloir regarder la ravissante vallée de la Cise, elle se rejeta promptement dans le fond de la calèche, et dit d'une voix qui en plein air paraissait d'une extrême faiblesse :

Als nächstes werden die Einleitungssätze mit dem TreeTagger getaggt. Da der TreeTagger als Input nur Dateien akzeptiert, in denen jedes Wort einzeln in einer Zeile steht, muss der auszuzeichnende Text erst noch dafür vorbereitet werden. Dazu wird der Text tokenisiert, d.h. in Einzelwörter zerlegt, und in eine temporäre Datei (ein Token pro Zeile) geschrieben, die im Anschluss daran durch den TreeTagger verarbeitet wird. Der dabei entstehende Output wird zurück in ein Array @output geleitet und wieder in die ursprüngliche Form zurückformatiert, mit dem einzigen Unterschied, dass die Einzelwörter nun mit Tags versehen sind.

Um die durchschnittliche Position des VC zu ermitteln, werden die Sätze, die sich wiederum in durch Kommata voneinander separierte Teilsätze gliedern, am Komma aufgetrennt und teilsatzweise in ein temporäres Array geschrieben. Jedes Element dieses Arrays wird daraufhin untersucht, ob eine konjugierte Verbform oder ein Infinitiv enthalten ist, d.h. jeder der Teilsätze wird vom letzten Satzelement ausgehend bis zurück zum Anfang des Satzes nach dem Muster / (VER:zeit) /²⁹, z.B. / (VER:impf) /, durchsucht. Wurde eine Verbform ermittelt, wird überprüft, ob diese in einer der zuvor erzeugten Verblisten³⁰, d.h. in `vd_konjug_o_pp.txt`, `vd_inf.txt` oder `vd_pp.txt`, enthalten ist. Stimmt die Verbform mit einer Form in der Liste überein, wird festgehalten, an welcher Position, d.h. in welchem Teilsatz vor der direkten Rede das Verb steht. Anschließend wird für alle redeeinleitenden Sätze der durchschnittliche Wortabstand des Redeverbs bis zum Beginn der direkten Rede berechnet. Auch der das Redeverb enthaltende Teilsatz wird in einem Array gespeichert (@vd_konstr_eingeleitet). Diese Informationen werden allesamt in der HTML-Datei `/output/dateiname/vd_eingeleitet_position_dateiname.html` festgehalten. Abbildung 3.4 zeigt eine solche Datei mit Angabe der Position der *verba communicandi*.

Die Zahl in Klammern zu Beginn jeder Zeile identifiziert den Abschnitt, in dem der einleitende Satz gefunden wurde. Ein Mausklick auf die als Link dargestellte Abschnittsnummer ermöglicht die Ansicht des vollständigen einleitenden Satzes inklusive der Tags, die bei Bedarf auch ausblendbar sind. Ein einleitender *discours citant* besteht aus ma-

²⁹ „zeit“ steht für das Kürzel der jeweiligen Tempusangabe, vgl. Tabelle 3.2, „Tagset für Französisch“.

³⁰ Vgl. Kapitel 3.3.4.

sue_mysteres_de_paris_1842.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Position des *verbum communicandi* in einleitenden Sätzen

Durchschnittliche Entfernung des *verbum communicandi* von der direkten Rede (in Teilsätzen (1-7)): 1,3

(x = konjugiertes VC in einleitendem (Teil-)Satz vor DR;

i = VD in Infinitivform; je kleiner die vorausgehende Zahl, desto näher steht das VD an der direkten Rede;

die Zahl in Klammern gibt die Abschnittnummer im Text wieder)

(386) 1x 2 3 4 5 6 7

(603) 1x 2 3 4 5 6 7

(630) 1 2 3 4 5 6 7

(850) 1x 2 3 4 5 6 7

(873) 1x 2 3 4 5 6 7

(1202) 1x 2 3 4 5 6 7

Abbildung 3.4: Auszug aus der Exportdatei `vd_eingeleitet_position_sue_mystères_de_paris_1842.html` (Screenshot), die über die Position des *verbum communicandi* in redeeinleitenden Sätzen informiert. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Die sich daran anschließende Tabelle bildet ab, innerhalb welchen Teilsatzes das Redeverb ermittelt wurde. Die Zahl in Klammern zu Beginn einer Dialogstelle gibt an, an welcher Stelle im Text sie sich befindet, bzw. macht deutlich, ob sich Dialogstellen in unmittelbarer Nähe zueinander befinden. Ein Klick auf die als Hyperlink dargestellte Abschnittnummer führt zur Anzeige des kompletten Einleitungssatzes. Das *verbum communicandi* ist jeweils durch Fettdruck und Großbuchstaben hervorgehoben. Untersucht wurden bis zu sieben sich vor der direkten Rede befindenden Teilsätze, dargestellt durch die Ziffern 1-7. Je kleiner der Zahlenwert, desto geringer ist die Entfernung des Redeverbs von der direkten Rede. Ein x hinter der Ziffer bedeutet, dass es sich um eine konjugierte Verbform handelt, während ein i einen Infinitiv anzeigt.

ximal sieben Teilsätzen³¹, dargestellt durch die Ziffern 1-7. Je kleiner der Zahlenwert, desto geringer ist auch die Entfernung von der direkten Rede, d.h. der Teilsatz mit der Ziffer 7 ist sieben Teilsätze von der direkten Rede entfernt, während sich Teilsatz 1 unmittelbar davor befindet. Ein „x“ hinter der Ziffer besagt, dass es sich bei dem identifizierten

³¹ Diese Regel ist arbiträr, beruht jedoch auf der Erkenntnis, dass kaum jemals ein VC im sechsten bzw. im siebten Teilsatz vor Beginn der direkten Rede vorkommt.

VC um eine konjugierte Form handelt, während ein „i“ ein VC in der Infinitivform darstellt.

Die abstrakte Darstellung gibt Hinweise darauf, welche Teilsätze hinsichtlich der automatischen Weiterverarbeitung problematisch und welche weniger kompliziert sein dürften. Unproblematisch für die Weiterverarbeitung sind all jene einleitenden Sätze, in denen jeweils nur ein konjugiertes *verbum communicandi* zu finden ist, das der direkten Rede vorangeht bzw. an vorletzter Stelle steht, wie z.B.

(2581) *Elle répondit posément, dans une colère froide : (2582) – C'est bon, en v'là assez.... Attends quinze jours, et je ne te gênerai plus, si c'est ça que tu demandes. Oui, dans quinze jours, j'aurai vingt et un ans, je filerai.*³²

Komplizierter wird es, wenn in mehreren Teilsätzen mehrere konjugierte *verba communicandi* anzutreffen sind, z.B. je ein *verbum communicandi* an Position eins und zwei:

(3840) *Mais elle reprit doucement, et avec un de ces regards qui implorant le mensonge .*³³

oder je ein VC an Position zwei und drei:

(4965) *Trop émue pour avoir parlé jusque-là, elle s's'écria, en s'asseyant à côté de Jacques et en remarquant sa pâleur .*³⁴

Dasselbe gilt für Einleitungssätze, in denen sich ein oder mehrere *verba communicandi* aus konjugierten Formen und auch Infinitivformen zusammensetzen, wie z.B. in

(3123) *dit la duègne, et elle ajouta entre ses dents, sans que Florine pût l'entendre .*³⁵

Aufgrund der Polysemie vieler *verba communicandi* ist leider nicht immer gewährleistet, dass ein durch das Programm identifiziertes VC auch tatsächlich eines ist:

(77) *Barbicane, ayant d'un geste rapide assuré son chapeau sur sa tête, continua son discours d'une voix calme .*³⁶

Steht weder ein „x“ noch ein „i“ hinter der Ziffer, wurde kein VC ermittelt. Das bedeutet, dass, anders als bei den Sätzen, die sich direkt an die direkte Rede anschließen, auch ein beliebiges anderes Verb die alleinige redeeinleitende Funktion übernehmen kann, wie in folgendem Beispiel:

(549) *Il rejoignit Rastignac, qu'il prit familièrement par le bras .*³⁷

Innerhalb des redeeinleitenden *discours citant* ist auch die Kombination aus einer beliebigen konjugierten Verbform in Verbindung mit einem VC in *gérondif*-Form recht häufig:

³² Émile Zola, *La terre* (1887).

³³ Gustave Flaubert, *L'éducation sentimentale* (1869).

³⁴ Eugène Sue, *Le juif errant* (1845).

³⁵ Eugène Sue, *Le juif errant* (1845).

³⁶ Jules Verne, *De la terre à la lune* (1865).

³⁷ Honoré de Balzac, *Le père Goriot* (1843).

(927) *L'abbé se leva, je lui pris le bras pour m'y appuyer en lui disant* :³⁸

Der einleitende *discours citant* kann jedoch auch ganz ohne VC auskommen, wie es das folgende Beispiel zeigt:

(16) *De là cette autre phrase* :³⁹

3.3.7 Suche nach Redebegleitung und Bestimmen der *verba communicandi*

An dieser Stelle im Programm wird auf das zu einem früheren Zeitpunkt⁴⁰ erzeugte Array @vd_dir_rede_satzanfang zurückgegriffen. Nach dem Taggen der Tokens in allen redebegleitenden Sätzen⁴¹ wird jedes einzelne Element – dabei handelt es sich um eine Dialogstelle, an deren ersten Position ein Dialogmarker (Bindestrich(e) oder Anführungszeichen) steht – in ganze Sätze zerlegt. Das Satzende gilt als erreicht, wenn ein Punkt, Ausrufe- oder Fragezeichen steht. Von diesen können auch mehrere aufeinanderfolgen. Ist nur ein einziger ganzer Satz ohne Komma enthalten, handelt es sich um einen unbegleiteten Turn, wie z.B.

(2118) – *Est-ce que le père de cet homme buvait ?*⁴²

Ein alleinstehender Turn liegt ebenfalls vor, wenn der einzelne Satz zwar ein oder mehrere Kommata enthält, aber kein Verb:

(362) – *C'est que, voyez-vous, j'adore le boudin chaud, quand il sort de la marmite. Je viendrai vous en chercher.* ⁴³

Diese unbegleiteten Turns werden zwischengespeichert und in eine Text- und HTML-Datei mit Namen turns_o_redebegleitung__\$dateiname_fuer_output.(txt|html) geschrieben.

Enthält eine Dialogstelle mehrere ganze Sätze, wird jeder einzelne ganze Satz jeweils durch Trennung an den Satzzeichen Komma und Semikolon in Teilsätze zerlegt. Für jeden dieser Teilsätze wird anschließend geprüft, ob die für einen redebegleitenden Satz typische Abfolge enthalten ist:

Satzzeichen - (Dativ-)Pronomen (fak.)⁴⁴ - konjugierte

³⁸ Honoré de Balzac, *Le lys dans la vallée* (1843).

³⁹ Honoré de Balzac, *Eugénie Grandet* (1843).

⁴⁰ Siehe Kapitel 3.3.3.

⁴¹ Beschrieben in Kapitel 3.3.5 und 3.3.6.

⁴² Émile Zola, *L'assommoir* (1877).

⁴³ Émile Zola, *Le ventre de Paris* (1873).

⁴⁴ (fak.) = Abkürzung für fakultativ, d.h. das Element kann, muss aber nicht vorkommen.

Verbform- konjugierte Verbform (fak.) - Klitikon (fak.) -
beliebig viele getaggte Wörter - Satzzeichen⁴⁵

Trifft dies zu, wird weiterhin geprüft, ob das Hauptverb der analytischen Verbkonstruktion oder die synthetische Verbform mit einem *verbum communicandi* aus der Liste `vd_pp.txt` übereinstimmt. Ist auch diese Bedingung erfüllt, werden sowohl das Hauptverb in dem Array `@einschub_nur_vd` als auch der komplette redebegleitende Satz in dem Array `@einschub_mit_vd` zur Weiterverarbeitung zwischengespeichert.

Der Teilsatz, auf den das Suchmuster passte, wird mit der Nummer, die zu Beginn jedes Elements steht, versehen und zur Weiterverarbeitung in das Array kopiert. Die auf diese Weise ermittelten redebegleitenden Sätze werden später weiter sortiert nach den innerhalb des redebegleitenden *discours citant* am häufigsten anzutreffenden Konstruktionstypen. Wird weder für den Einzelsatz noch die Teilsätze ein *verbum communicandi* nachgewiesen, wird angenommen, dass keine Redebegleitung vorliegt, und der Satz zu den unbegleiteten Turns in dem Array `@einschub_o_rb` hinzugefügt.

3.3.8 Ermitteln der verschiedenen Konstruktionstypen

Sowohl im redeeinleitenden als auch im redebegleitenden *discours citant* kehren bestimmte Formulierungsmuster immer wieder. Odile Halmøy (1983) beschreibt die verschiedenen Konstruktionen, die für den redebegleitenden *discours citant* typisch sind.⁴⁶ Diese Ausführungen stellten für die Programmierung von DCAnalyse, genauer für die Formulierung der Regulären Ausdrücke, eine wertvolle Hilfe dar.

Es ergeben sich die folgenden, bereits in Kapitel 2.2.3 vorgestellten Konstruktionstypen, nach denen – von positionsbedingten (Redeeinleitung bzw. Redebegleitung) Abweichungen abgesehen – die vom Programm ermittelten redeeinleitenden und redebegleitenden Sätze durchsucht werden:

- **T1** / BASISTYP:
« ... », *dit-il*.
- **T2a** / KONSTRUKTION MIT ADJEKTIV ODER *participe passé*:
« ... », *répliqua-t-elle stupéfaite*.
- **T2b** / KONSTRUKTION MIT ADJEKTIV ODER *participe présent*:
« ... », *dit la jeune fille trouvant enfin la force de parler dans son émotion*.
- **T2c** / RELATIVSATZ-KONSTRUKTION:
« ... », *dit une voix qui ne lui était pas inconnue*.

⁴⁵ Ein Beispiel für einen solchen Regulären Ausdruck ist in Tabelle 3.3 dargestellt.

⁴⁶ Eine ausführlichere Beschreibung erfolgt in Kapitel 2.2.3.

- **T3a** / ADVERBIALKONSTRUKTION:
« ... », *s'écria vivement le jeune homme.*
- **T3b** / *gérondif*-KONSTRUKTION:
« ... », *remarqua-t-il en souriant.*
- **T3c** / PRÄPOSITIONALSYNTAGMA:
« ... », *répondit-elle avec quelque impatience.*
- **T3e** / KONSTRUKTION MIT DATIV-ERGÄNZUNG:
« ... », *murmura-t-elle à son mari.*
- **Tor** / KONSTRUKTION MIT OREILLE:
« ... », *dit l'oncle à l'oreille du neveu.*
- **Tvoix** / KONSTRUKTION MIT VOIX/TON/AIR:
« ... » ? *demanda-t-elle d'une voix altérée.*
- **TKon** / KONSTRUKTION MIT KONJUNKTION:
« ... », *dit César quand il la vit endormie.*

Suche nach Konstruktionstypen innerhalb der Redebegleitung

Das Weiterverarbeiten der redebegleitenden Sätze erfolgt auf der Grundlage des Arrays `@einschub_mit_vd`⁴⁷, in dem alle redebegleitenden Sätze enthalten sind. Bevor mit dem Ermitteln der Konstruktionstypen begonnen werden kann, müssen zunächst noch durch den TreeTagger herbeigeführte Falschauszeichnungen⁴⁸ korrigiert werden, d.h. fehlerhafte Tags werden, sofern das möglich und für die weitere Analyse von Bedeutung ist, verbessert. An den Ursachen für Falschauszeichnungen ändert dies jedoch nichts, so dass gerade bei der Bearbeitung von neuem Textmaterial neue Falschauszeichnungen hinzukommen können, die von der Software nicht berücksichtigt werden können.

In einem weiteren Schritt werden die Regulären Ausdrücke für alle Konstruktionstypen formuliert. Wie diese aussehen, wird für den Basis-Typ («*dit-il*», Typ 1), die Adverbialkonstruktion (Typ 3a), die *gérondif*- (Typ 3b) und die Dativkonstruktion (Typ 3e) anhand der Tabellen 3.4 bis 3.7 exemplarisch gezeigt.

Als nächstes wird das Array `@einschub_mit_vd` auf Übereinstimmung mit einem dieser Ausdrücke hin überprüft. Gibt es einen Treffer, wird dieser in einem weiteren, dem Konstruktionstyp entsprechenden Array gespeichert. Analog dazu werden auch alle Elemente, auf die keiner der Ausdrücke passt, in einem „Reste-Array“ gesammelt. Wurden alle redebegleitenden Sätze auf mögliche Konstruktionen hin überprüft, wird die Datei `vd_konstruktionen_eingeschoben_$dateiname` erzeugt, die wahlweise als reine

⁴⁷ Details zur Erstellung in Kapitel 3.3.7.

⁴⁸ Beispiele siehe Kapitel 3.4.2.

Text- oder als HTML-Datei konsultierbar ist. Außerdem wird berechnet, wie oft jeder Konstruktionstyp vorkommt und welchen prozentualen Anteil dieser an der gesamten Redebegleitung hat, was wiederum die Grundlage für die in Kapitel 4.2.2 enthaltenen Schaubilder bildet.

Tag	Bedeutung	Tag	Bedeutung
PRO:IND	indefinite pronoun	VER:impf	verb imperfect
PRO	pronoun	VER:simp	verb simple past
PUN:cit	punctuation citation	SENT	sentence tag
VER:subi	verb subjunctive imperfect	VER:cond	verb conditional
NUM	numeral	NOM	noun
VER:pper	verb past participle	VER:futu	verb futur
PRO:PER	personal pronoun	VER:impe	verb imperative
NAM	proper name	PRO:POS	possessive pronoun
PRP	preposition	ADV	adverb
PRP:det	preposition plus article	INT	interjection
DET:ART	article	DET:POS	possessive pronoun
ABR	abbreviation	PRO:REL	relative pronoun
PRO:DEM	demonstrative pronoun	ADJ	adjective
VER:pres	verb present	VER:infi	verb infinitive
KON	conjunction	SYM	symbol
VER:subp	verb subjunctive present	PUN	punctuation
VER:ppre	verb present participle		

Tabelle 3.2: TreeTagger: Tagset-Schlüssel für Französisch. Quelle: <http://www.ims.uni-suttgart.de/~schmid/french-tagset.html>.

Regulärer Ausdruck für die Suche nach redebegleitendem <i>discours citant</i>		
/(\$pun(\$seluileur)?(\$verb)?(\$verb)(\$pro_per_klitisch)?(\$wort_tagged)*\$pun)/		
findet Konstruktionen wie		
[, (PUN)][RÉPONDIT (VER:simp)][-il (PRO:PER)][vivement (ADV)][. (SENT)]		
Teilausdruck	RegEx	Bedeutung
\$pun	(\[[; , ! \ ? \ . :] \ \ ((PUN SENT)\)\)\]{1,3}	1-3 Satzzeichen
\$seluileur	\[(s[\ ' e] lui leur) \ \ (PRO:PER)\ \]	Dativpronomen
?	?	0-1 Vorkommen
\$verb	\[[A-Za-z-âäëèéêçôîââëèéèîîôç]* \ \ (VER: [a-z]{4}\)\ \]	Verbform
\$pro_per_klitisch	\[-?t?-(je tu ils? elles? nous vous) \ \ (PRO:PER)\ \]	Klitikon
\$wort_tagged	\[[0-9A-Za-z-âäëèéèîîôùçç\-\ ' : \ (\)]*\ \]	getaggte Wortform
*	*	beliebig viele Vorkommen („Wildcard“)

Tabelle 3.3: Regulärer Ausdruck für die Suche nach einem beliebigen redebegleitenden *discours citant*. In der zweiten Zeile ist der Ausdruck abgebildet, wie er auch im Perl-Skript vorkommt, jedoch verkürzt dargestellt durch eine Abfolge von Variablen. In der vierten Zeile folgt ein getaggtter Beispielsatz, auf den der Ausdruck passt. In den unteren zwei Dritteln der Tabelle werden die in der zweiten Zeile verwendeten Variablen aufgelöst in die eigentlichen Regulären Ausdrücke.

Regulärer Ausdruck für die Suche nach T1 («dit-il»)		
/^\$el_zahl\$pun(\$pro_per_reflexiv)?\$verb\$pro_per_klitisch\$pun/		
findet Sätze wie		
(1234)[, (PUN)][[REPRIT (VER:simp)][-elle (PRO:PER)]]. (SENT)] ⁴⁹		
Teilausdruck	RegEx	Bedeutung
^	^	Zeilenanfang
\$el_zahl	\\(\\d*)	Elementzahl
\$pun	(\\[[; , ! \\? \\ . :] \\ \\((PUN SENT)\\)\\)\\){1,3}	Satzzeichen
\$pro_per_reflexiv	\\[[ms] [e\\'] \\ \\(PRO:PER)\\]	Reflexivpronomen
?	?	0-1 Vorkommen
\$verb	\\[[A-Za-z-äåêëëçôîâäéëëîïöç]* \\ (VER: [a-z]{4})\\]	Verbform
\$pro_per_klitisch	\\[-?t?-(je tu ils? elles? nous vous) \\ (PRO:PER)\\]	Klitikon
\$pun	(\\[[; , ! \\? \\ . :] \\ \\((PUN SENT)\\)\\)\\){1,3}	Satzzeichen

Tabelle 3.4: Regulärer Ausdruck für die Suche nach T1 («dit-il»). In der zweiten Zeile ist der Ausdruck abgebildet, wie er auch im Perl-Skript vorkommt, jedoch verkürzt dargestellt durch eine Abfolge von Variablen. In der vierten Zeile folgt ein getaggtter Beispielsatz, auf den der Ausdruck passt. In den unteren zwei Dritteln der Tabelle werden die in der zweiten Zeile verwendeten Variablen aufgelöst in die eigentlichen Regulären Ausdrücke.

Regulärer Ausdruck für die Suche nach T3a (Adverbialkonstruktion)		
/^\$el_zahl\$pun(\$pro_per_reflexiv)?\$verb\$pro_per_klitisch\$adv(\$adv)?/		
findet Sätze wie		
(532)[? (SENT)][[DIT (VER:pres)][-il (PRO:PER)][[brusquement (ADV)]]. (SENT)]		
Teilausdruck	RegEx	Bedeutung
^	^	Zeilenanfang
\$el_zahl	\\(\\d*)	Elementzahl
\$pun	(\\[[; , ! \\? \\ . :] \\ \\((PUN SENT)\\)\\)\\){1,3}	Satzzeichen
\$pro_per_reflexiv	\\[[ms] [e\\'] \\ \\(PRO:PER)\\]	Reflexivpronomen
?	?	0-1 Vorkommen
\$verb	\\[[A-Za-z-äåêëëçôîâäéëëîïöç]* \\ (VER: [a-z]{4})\\]	Verbform
\$pro_per_klitisch	\\[-?t?-(je tu ils? elles? nous vous) \\ (PRO:PER)\\]	Klitikon
\$adv	\\[[a-z-âäéëëîïöçç]* \\ (ADV)\\]	Adverb

Tabelle 3.5: Regulärer Ausdruck für die Suche nach T3a (Adverbialkonstruktion). In der zweiten Zeile ist der Ausdruck abgebildet, wie er auch im Perl-Skript vorkommt, jedoch verkürzt dargestellt durch eine Abfolge von Variablen. In der vierten Zeile folgt ein getaggtter Beispielsatz, auf den der Ausdruck passt. In den zwei unteren Dritteln der Tabelle werden die in der zweiten Zeile verwendeten Variablen aufgelöst in die eigentlichen Regulären Ausdrücke. Die hier gezeigt Kombination aus Regulären Ausdrücken findet nur redebegleitende Sätze, in denen die Bezugnahme auf das sprechende Subjekt durch ein Pronomen markiert wird. Nach anderen möglichen Kombinationen (z.B. «...? dit simplement Mme Desforges.» (aus: Émile Zola, *Au bonheur des dames* (1883), Satz (2191))) wird natürlich ebenfalls gesucht.

3.3.9 Export der Ergebnisse

DCAnalyse erzeugt umfangreiches Datenmaterial, das entweder zur direkten Auswertung verwendet kann oder sich zur Weiterverarbeitung eignet. Für die Ausgabe werden Klartextdateien im Text-, HTML- oder CSV⁵⁰-Format erzeugt, die nach dem ProgrammDurchlauf im eigens dafür angelegten Verzeichnis `output` gespeichert werden. Die Textdateien sind mit Hilfe jedes beliebigen Texteditors lesbar, während die `.csv`-Dateien mit allen gängigen Tabellenkalkulationsprogrammen (z.B. OpenOffice Calc, Microsoft Excel, iWorks Numbers u.v.a.) geöffnet und weiterverarbeitet werden können. Da in den Datenzellen auch Kommata vorkommen können, jedoch keine Semikola, wurde als Trennzeichen das Semikolon anstelle des normalerweise üblichen Kommas gewählt. Dies muss beim Öffnen der Datei mit dem jeweiligen Tabellenkalkulationsprogramm berücksichtigt werden.

Auf der obersten Ebene dieses Verzeichnisses `/output` befinden sich drei `.csv`-Dateien, die grundlegende Informationen zum Inhalt der Datei und dem *discours citant* im jeweils bearbeiteten Text liefern:

- `datenblatt_allgemein.csv`
- `datenblatt_vd_redebegleitung.csv`
- `datenblatt_vd_redeeinleitung.csv`

Die Datei `datenblatt_allgemein.csv` gibt Auskunft über die Gesamtwortzahl, den Anteil der direkten Rede am Text, die Anzahl der Dialogstellen sowie die Frequenz der ermittelten VC-Typen und -Tokens, jeweils unterschieden nach Redeeinleitung und Redebegleitung. Die Datei `datenblatt_vd_redebegleitung.csv` enthält Angaben zu den redebegleitenden Sätzen und listet die Dialogstellen, die zu Beginn einer Zeile gefunden wurden, die darin enthaltenen VC-Typen und -Tokens sowie die Häufigkeit der verschiedenen Konstruktionen in den redebegleitenden Sätzen auf. Analog zur vorherigen Datei finden sich in `datenblatt_vd_redeeinleitung.csv` die gleichen Informationen zu den redeeinleitenden Sätzen. Unterhalb des Verzeichnisses `output/` wird für jede bearbeitete Datei ein eigenes Verzeichnis angelegt. Heißt die Datei `autor1867.txt`, so wird ein Verzeichnis namens `autor1867/` erstellt, in welchem weitere Exportdateien im Text- als auch im HTML-Format gesammelt werden. Die Darstellung der Inhalte mittels HTML-Dateien wurde als zusätzliches Ausgabeformat gewählt, um zum einen durch Änderungen am Layout die Lesbarkeit gegenüber reinen Textdateien zu verbessern und zum anderen das zügige Navigieren innerhalb der Export-Dateien durch Hyperlinks zu

⁵⁰ Abkürzung für engl. *Comma Separated Values*. Dieses Dateiformat dient der Speicherung einfach strukturierter Daten, z.B. wie im vorliegenden Fall zum Erzeugen einer Tabellenstruktur.

voltaire_candide1759.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Ausgabe aller Turns mit Redeeinleitung

Anzahl: 17 Turns mit Redeeinleitung

(37) Cet orateur, le regardant de travers, lui dit: (38) --Que venez-vous faire ici? y êtes-vous pour la bonne cause?

(50) Aussitôt Candide le mène dans l'étable de l'anabaptiste, où il lui fit, manger un peu de pain; et quand Pangloss fut refait: (51) --Eh bien! lui dit-il, Cunégonde?

(71) Le matelot disait en sifflant et en jurant: (72) --Il y aura quelque chose à gagner ici.

(75) Pangloss le tirait cependant par la manche: (76) --Mon ami, lui disait-il, cela n'est pas bien, vous manquez à la raison universelle; vous prenez mal votre temps.

(78) Il disait à Pangloss: (79) --Hélas! procure-moi un peu de vin et d'huile; je me meurs.

(321) Le vieillard rougit encore: (322) --Est-ce qu'il peut y avoir deux religions? dit-il. Nous avons, je crois, la religion de tout le monde; nous adorons Dieu du soir jusqu'au matin.

(352) Candide dit à Cacambo: (353) --Mon ami, vous voyez comme les richesses de ce monde sont périssables; il n'y a rien de solide que la vertu et le bonheur de revoir mademoiselle Cunégonde.

(399) En un mot, j'en ai tant vu et tant éprouvé que je suis manichéen: (400) --Il y a pourtant du bon, répliquait Candide.

Abbildung 3.5: Beispiel für einen Auszug aus der von DCAnalyse erzeugten HTML-Exportdatei `dir_rede_eingeleitet_voltaire_candide_1759.html` (Screenshot), die alle redeeinleitenden Sätze und die nachfolgende direkte Rede aufführt. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Jeder Dialogstelle geht eine Ziffer in Klammern voraus, anhand derer zu erkennen ist, an welcher Stelle im Text sie sich befindet und welche Dialogstellen direkt aufeinanderfolgen.

ermöglichen. So kann der Betrachter per Link zu weiteren HTML-Export-Dateien springen und zwischen der Ansicht mit bzw. ohne TreeTagger-Tags hin- und herwechseln. Anhand der Textdatei `datei.txt` lässt sich überprüfen, ob der Inhalt der Input-Datei korrekt eingelesen wurde. Treten hier bereits Fehler auf, muss der Benutzer die Eingabedatei vor der Bearbeitung mit DCAnalyse erst noch manuell anpassen. Die Datei `dir_rede_alle.(txt|html)` enthält alle Turns, die entweder eingeleitet, begleitet oder

flaubert_madame_bovary1857.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Ausgabe aller Turns (mit oder ohne Redebegleitung)

Anzahl: 966 Turns mit oder ohne Redebegleitung

- (8) -- Levez-vous, dit le professeur.
(11) -- Débarrassez-vous donc de votre casque, dit le professeur, qui était un homme d'esprit.
(13) -- Levez-vous, reprit le professeur, et dites-moi votre nom.
(15) -- Répétez !
(17) -- Plus haut ! cria le maître, plus haut !
(21) -- Que cherchez-vous ? demanda le professeur.
(22) -- Ma cas..., fit timidement le nouveau , promenant autour de lui des regards inquiets.
(23) -- Cinq cents vers à toute la classe ! exclamé d'une voix furieuse, arrêta, comme le Quos ego , une bourrasque nouvelle.
(24) -- Restez donc tranquilles ! continuait le professeur indigné, et s'essuyant le front avec son mouchoir qu'il venait de prendre dans sa toque : Quant à vous, le nouveau , vous me copierez vingt fois le verbe ridiculus sum .
(60) -- Etes-vous le médecin ? demanda l'enfant.

Abbildung 3.6: Beispiel für einen Auszug aus der von DCAnalyse erzeugten HTML-Exportdatei `dir_rede_satzanfang_flaubert_bovary_1857.html` (Screenshot), die alle Turns mit Redebegleitung auflistet. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Jeder Dialogstelle geht eine Ziffer in Klammern voraus, anhand derer zu erkennen ist, an welcher Stelle im Text sie sich befindet und welche Dialogstellen direkt aufeinanderfolgen.

alleinstehend sind. In `dir_rede_eingeleitet.(txt|html)` sind alle Turns aufgelistet, denen eine Redeeinleitung vorausgeht. `dir_rede_satzanfang.(txt|html)` enthält alle Turns, die zu Beginn eines Abschnittes stehen und denen in der Regel ein redebegleitender Satz folgt. Wie in Kapitel 3.3.3 bereits erwähnt wurde, sind darin auch noch alleinstehende Turns enthalten.

Im Folgenden werden zur besseren Anschauung exemplarisch einige Dateien gezeigt (siehe Abbildungen 3.4 bis 3.10). Die redeeinleitenden Sätze werden in der Da-

tei vd_konstruktionen_eingeleitet_m_tags.(txt|html) sortiert nach den vorkommenden Konstruktionen gespeichert. Die *verba communicandi* werden jeweils durch Großbuchstaben und Fettdruck graphisch hervorgehoben.

zola_germinal1885.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Übersicht über die Konstruktionstypen in redeeinleitenden Sätzen:

[TreeTagger-Tags ausblenden](#)

Typ 1 ("il répondit"): 56 von 115 (48,7 %)

(14) [, (PUN)][Étienne (NAM)][**DEMANDA** (VER:simp)][, (PUN)]

(53) [, (PUN)][quand (KON)][il (PRO:PER)][put (VER:simp)][**PARLER** (VER:infi)][: (PUN)]

(68) [Il (PRO:PER)][**DEMANDA** (VER:simp)][: (PUN)]

(73) [, (PUN)][il (PRO:PER)][**DIT** (VER:pres)][, (PUN)]

(126) [, (PUN)][Maheu (NAM)][parut (VER:simp)][tout (ADV)][d' (PRP)][un (DET:ART)][coup (NOM)][les (PRO:PER)][**ENTENDRE** (VER:infi)][, (PUN)]

(146) [C' (PRO:DEM)][était (VER:impf)][la (DET:ART)][Maheude (NOM)][qui (PRO:REL)][**CRIAIT** (VER:impf)][: (PUN)]

Abbildung 3.7: Beispiel für einen Auszug aus der von DCAnalyse erzeugten HTML-Exportdatei vd_konstruktionen_eingeleitet_zola_germinal_1885.html (Screenshot), die die redeeinleitenden Sätze unter Hervorhebung der *verba communicandi* abbildet. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Jeder Dialogstelle geht eine Ziffer in Klammern voraus, anhand derer zu erkennen ist, an welcher Stelle im Text sie sich befindet und welche Dialogstellen direkt aufeinanderfolgen. Das *verbum communicandi* ist durch Fettdruck und Großbuchstaben hervorgehoben. Die TreeTagger-Tags sind bei Bedarf mittels Hyperlink (TreeTagger-Tags ausblenden) aus- und wieder einblendbar.

Die Anzeige der TreeTagger-Tags kann in der HTML-Version durch Anklicken des Links TreeTagger-Tags ausblenden unterdrückt werden. Die Anzeige wechselt in diesem Fall auf die Datei vd_konstruktionen_eingeleitet_o_tags.html. Ein Hin- und Herwechseln zwischen den Dateien ist durch entsprechende Hyperlinks gewährleistet.

Eine weitere Datei namens `vd_konstruktionen_eingeschoben.(txt|html)` führt die *verba communicandi* in redebegleitenden Sätzen (mit TreeTagger-Tags) auf. Die Tags können durch Anklicken des Hyperlinks `Blende TreeTagger-Tags` aus unterdrückt werden, woraufhin die Datei `vd_konstruktionen_eingeschoben_m_tags.html` angezeigt wird. Die Datei `turns_o_redebegleitung.(txt|html)` listet die alleinstehenden Turns, bei denen weder ein redeeinleitender noch ein redebegleitender *discours citant* steht, auf.

Eine detaillierte tabellarische Auflistung der *verba communicandi* in redeeinleitenden und redebegleitenden Sätzen ist der Datei `frequ_vd_uebersicht.html` zu entnehmen. Die Sortierung erfolgt nach der Häufigkeit ihrer Vorkommen. Da eine übersichtliche tabellarische Darstellung im Textformat nicht möglich ist, wird diese Datei nur im HTML-Format erzeugt.

proust_du_cote_de_chez_swann_1913.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Übersicht über die Konstruktionstypen innerhalb der redebegleitenden Sätze

[Typ 1a \("dit-il"\)](#)

[Typ 1b \("dit la comtesse"\)](#)

[Typ 1c \("dit Julie"\)](#)

[Typ 2a \("s'exclama Zazie furieuse / dit Gabriel conciliant"\)](#)

[Typ 2b \("répond Pierre tombant dans le piège"\)](#)

[Typ 2c \("dit-il" + phrase relative\)](#)

[Typ 3a \("dit-il" + adverbe\)](#)

[Typ 3b \("dit-il" + *gérondif*\)](#)

[Typ 3c \("dit-il" + syntagme prépositionnel\)](#)

[Typ 3d \("dit-il" + sans + infinitif\)](#)

[Typ 'oreille' \("dit-il à l'oreille de"\)](#)

[Typ 3e1 \("dit-il" + complément datif\)](#)

[Typ 3e2 \("dit-il" + complément datif\)](#)

[Typ 3e3 \("dit-il" + complément datif\)](#)

[Typ 3e5 \("dit-il" + complément datif\)](#)

[Typ 'voix' \("dit-il" + voix/ton/air\)](#)

[Typ "dit-il" + Konjunktion](#)

[Typ "faire" + Infinitiv](#)

[Rest \(keine Zuordnung möglich\)](#)

Typ 1a ("dit-il"): 20 von 273 (7,3 %)

[Blende TreeTagger-Tags aus](#)

(252)[!(SENT)][AJOUTAIT (VER:impf)][-il (PRO:PER)][, (PUN)]

(285)[, (PUN)][DIT (VER:pres)][-elle (PRO:PER)][, (PUN)]

(290)[, (PUN)][DIT (VER:pres)][-elle (PRO:PER)][, (PUN)]

(300)[, (PUN)][AJOUTA (VER:simp)][-t-il (PRO:PER)][, (PUN)]

Abbildung 3.8: Beispiel für einen Auszug aus der von DCAnalyse erzeugten Exportdatei `vd_konstruktionen_eingeschoben_proust_du_cote_de_chez_swann_1913.html` (Screenshot). Die redebegleitenden Sätze werden sortiert nach den für die Redebegleitung möglichen Konstruktionstypen angezeigt. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Die Hyperlinks direkt darunter führen zu der entsprechenden Konstruktion. Jeder Dialogstelle geht eine Ziffer in Klammern voraus, anhand derer zu erkennen ist, an welcher Stelle im Text sie sich befindet und welche Dialogstellen direkt aufeinanderfolgen. Die TreeTagger-Tags sind bei Bedarf mittels Hyperlink (Blende TreeTagger Tags aus) aus- und wieder einblendbar.

balzac_file_aux_yeux_d_or1835.html

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Turns ohne Einleitung bzw. redebegleitenden Satz

Anzahl: 73

[Blende TreeTagger-Tags ein](#)

(29) -- Il est très-fort . est l' immense éloge décerné à ceux qui sont arrivés , quibuscumque viis , à la politique , à une femme ou à une fortune . Parmi eux , se rencontrent certains jeunes gens qui jouent ce rôle en le commençant avec des dettes ; et naturellement , ils sont plus dangereux que ceux qui le jouent sans avoir un sou .

(34) -- Demandez à de Marsay , et vous verrez !

(42) -- J' allais te faire la même question .

(43) -- Une intrigue .

(44) -- Une intrigue ?

(45) -- Bah !

(46) -- e puis bien te dire cela à toi , sans compromettre ma passion . Puis une femme qui vient le dimanche aux Tuileries n' a pas de valeur , aristocratiquement parlant .

(47) -- Ha ! ha !

(48) -- Tais -toi donc , ou je ne te dis plus rien . Tu ris trop haut , tu vas faire croire que nous avons trop déjeuné . Jeudi dernier , ici , sur la terrasse des Feuillants , je me promenais sans penser à rien du tout . Mais en arrivant à la grille de la rue de Castiglione par laquelle je comptais m' en aller , je me trouve nez à nez avec une femme , ou plutôt avec une jeune personne qui , si elle ne m' a pas sauté au cou , fut arrêtée , je crois , moins par le respect humain que par un le ces étonnements profonds qui coupent bras et jambes , descendent le long de l' épine dorsale et s' arrêtent dans la plante des pieds pour vous attacher au sol . J' ai souvent produit des effets de ce genre , espèce de magnétisme animal qui devient très-puissant lorsque les rapports sont respectivement crochus .

Abbildung 3.9: Beispiel für einen Auszug aus der von DCAnalyse erzeugten Exportdatei turns_o_redebegleitung_balzac_fille_aux_yeux_d_or_1835.html (Screenshot). Auflistung der Turns, die weder durch einen *discours citant* eingeleitet noch begleitet werden. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Jeder Dialogstelle geht eine Ziffer in Klammern voraus, anhand derer zu erkennen ist, an welcher Stelle im Text sie sich befindet und welche Dialogstellen direkt aufeinanderfolgen. Die TreeTagger-Tags sind bei Bedarf mittels Hyperlink (Blende TreeTagger Tags ein) ein- und wieder ausblendbar.

hugo_miserables_complet_1862.txt

Navigation zu weiteren Ergebnisseiten:

[Übersicht über alle Turns](#)

[Übersicht über alle Turns zu Beginn einer Zeile](#)

[Übersicht über alle durch ":" eingeleiteten Turns](#)

[Frequenz der *verba communicandi*](#)

[Übersicht über die *verba communicandi* in redeeinleitenden Sätzen](#)

[Übersicht über die *verba communicandi* in redebegleitenden Sätzen](#)

[Position des *verbum communicandi* in einleitenden Sätzen](#)

[Turns ohne Einleitung bzw. redebegleitenden Satz](#)

[Überblick über die Konstruktionen *verbum communicandi* + *gérondif*](#)

Frequenz der *verba communicandi*

Gesamtwortzahl für hugo_miserables_complet_1862.txt: 557192

VC = verba communicandi

RE = Redeeinleitung

RB = Redebegleitung (eingeschobener Satz)

VC-Types (ges.)		VC-Tokens (ges.)		VC-Types (RE)		VC-Tokens (RE)		VC-Types (RB)		VC-Tokens (RB)	
VC Types (91)	Σ	VC Tokens (2145)	Σ	VD Types (76)	Σ	VD Tokens (616)	Σ	VD Types (48)	Σ	VD Tokens (1529)	Σ
accepter	1	accepta	1	accepter	1	accepta	1	achever	1	achève	1
achever	3	acheva	1	achever	2	acheva	1	ajouter	26	ajouta	24
admirer	1	achevait	1	admirer	1	achevait	1	amuser	2	ajoutaient	1
adresser	4	achève	1	adresser	4	admirait	1	appeler	1	ajoutait	1
ajouter	47	admirait	1	ajouter	21	adressa	4	balbutier	2	amuse	1
amuser	2	adressa	4	annoncer	1	ajouta	17	bégayer	1	amusent	1
annoncer	1	ajouta	41	appeler	3	ajoutaient	1	brailler	1	appelle	1
appeler	4	ajoutaient	2	appuyer	2	ajoutait	2	commander	1	balbutia	2
appuyer	2	ajoutait	3	articuler	1	ajouter	1	commencer	1	bégaya	1
articuler	1	ajouter	1	assurer	1	annonçait	1	conter	1	braille	1
assurer	1	amuse	1	balbutier	4	appelait	1	continuer	7	commandé, ont	1
balbutier	6	amusent	1	bégayer	3	appeler	1	crier	76	commença	1
bégayer	4	annonçait	1	chanter	1	appelle	1	décider	1	conte	1
brailler	1	appelait	1	commencer	5	appuya	1	déclarer	1	continua	7
chanter	1	appeler	1	congédier	1	appuyait	1	demander	80	cria	70
commander	1	appelle	2	conseiller	1	articulée	1	dire	720	criait	4
commencer	6	appuya	1	constater	2	assura	1	écrier	87	crie	1
congédier	1	appuyait	1	continuer	25	balbutia	3	exclamer	1	crièrent	1
conseiller	1	articulée	1	crier	37	balbutier	1	expliquer	1	décidé, ont	1
constater	2	assura	1	déclarer	1	bégaya	2	faire	81	déclara	1
conter	1	balbutia	5	demander	30	bégayer	1	fredonner	1	demanda	78

Abbildung 3.10: Beispiel für einen Auszug aus der von DCAnalyse erzeugten Exportdatei frequ_vd_uebersicht.html (Screenshot). Tabellarische Darstellung der Frequenz der *verba communicandi*. Die Hyperlinks in der oberen Bildhälfte ermöglichen die schnelle Navigation zu anderen Exportdateien. Die sechs Spalten bilden die VC-Types bzw. VC-Tokens jeweils für den gesamten *discours citant* (Redeeinleitung plus Redebegleitung) sowie nur für die Redeeinleitung und nur für die Redebegleitung unter Angabe der Anzahl der Types bzw. der Tokens ab.

3.4 Hürden beim Programmieren

Wie eingangs bereits erwähnt wurde, trägt die Qualität des E-Textes bzw. dessen Formatierung maßgeblich zu seiner automatisierten Verarbeitbarkeit bei. Mit vielen E-Texten lässt sich problemlos arbeiten. Dennoch enthalten einige Dateien problematische Formatierungen, durch die so mancher E-Text für eine Analyse unbrauchbar ist. Häufiger auftretende Formatierungsfehler ließen sich bei der Entwicklung des Programms berücksichtigen, indem der Fehler vor Beginn der eigentlichen Analyse korrigiert wurde. Jedoch ist es unmöglich, alle Abweichungen vorherzusehen, so dass auf die Analyse einzelner Texte im schlimmsten Falle ganz verzichtet werden muss. Eine weitere Fehlerquelle ergibt sich beim Taggen der Redeeinleitung und -begleitung. Die Zuweisung falscher Wortarten verhindert stellenweise das korrekte Arbeiten von DCAnalyse. Bei systematisch wiederkehrenden Falschauszeichnungen konnte dies bei der Programmierung berücksichtigt werden. Wo dies nicht möglich war, werden mitunter einzelne Fundstellen übergangen, die bei korrektem Tagging problemlos ermittelt werden.

3.4.1 Mangelhafte Textformatierung

In den beiden im Rahmen des *Project Gutenberg* digitalisierten Romanen *Pot-bouille* (1882) von Émile Zola und *Vingt-mille lieues sous les mers* (1865) von Jules Verne sind in der Textversion, die offiziell mit dem Zeichensatz ISO 8859-1 kodiert ist, keine Akzentzeichen enthalten (Beispiel siehe Abbildung 3.11). Vermutlich sind diese im Zuge der

Le pere, surpris, eut la naivete d'ajouter:

— Ah! tu crois?

Hortense et Berthe etaient devenues sourdes. Elles eurent pourtant un faible sourire, en affectant de s'occuper de leurs chaussures, qui etaient dans un pitoyable etat. Pour faire diversion, madame Josserand chercha une autre querelle a M. Josserand: elle le pria de remporter son journal chaque matin, de ne pas le laisser trainer tout un jour dans l'appartement, comme la veille par exemple; justement un numero ou il y avait un proces abominable, que ses filles auraient pu lire. Elle reconnaissait bien la son peu de moralite.

Abbildung 3.11: Mangelhafte Formatierung der Textdatei *Pot-bouille* (1882) von Émile Zola. Die Akzentzeichen fehlen, weshalb der Roman für die Analyse nicht zu gebrauchen ist und nicht in das Arbeitskorpus aufgenommen wurde.

Texterkennung abhandengekommen. Die Texte sind ohne Akzentzeichen für die Benutzung mit DCAnalyse unbrauchbar, da *verba communicandi* wie *s'écrier* oder *répliquer* nicht gefunden werden. Glücklicherweise ist *Vingt-mille lieues sous les mers* zusätzlich

im HTML-Format verfügbar, so dass der Roman dennoch in das Korpus aufgenommen werden konnte.

Bei reinen Textdateien kommt es immer wieder vor, dass um der Lesbarkeit im Anzeigemedium willen ein Zeilenumbruch nach einer bestimmten Zeilenlänge erfolgt (Beispiel siehe Abbildung 3.12). Was für den Betrachter am Bildschirm angenehm ist, wird für das Programm zum Stolperstein und muss für die die korrekte Verarbeitung erst wieder rückgängig gemacht werden, da das Programm den Zeilenumbruch als Beginn eines neuen Abschnitts interpretiert, obwohl das Satzende noch gar nicht erreicht ist.

```
Un dimanche, pendant ma lecture au jardin, je fus dérangé par Swann  
qui venait voir mes parents.  
  
--Qu'est-ce que vous lisez, on peut regarder? Tiens, du Bergotte? Qui  
donc vous a indiqué ses ouvrages?» Je lui dis que c'était Bloch.  
  
--Ah! oui, ce garçon que j'ai vu une fois ici, qui ressemble tellement  
au portrait de Mahomet II par Bellini. Oh! c'est frappant, il a les  
mêmes sourcils circonflexes, le même nez recourbé, les mêmes pommettes  
saillantes. Quand il aura une barbiche ce sera la même personne. En  
tout cas il a du goût, car Bergotte est un charmant esprit.» Et voyant  
combien j'avais l'air d'admirer Bergotte, Swann qui ne parlait jamais  
des gens qu'il connaissait fit, par bonté, une exception et me dit:
```

Abbildung 3.12: Mangelhafte Textformatierung in *Du côté de chez Swann* (1913) von Marcel Proust. Nach einer bestimmten Anzahl von Zeichen erfolgt ein Zeilenumbruch, der dem Leser die Lektüre am Bildschirm erleichtern soll. Die Software jedoch interpretiert jeden dieser Umbrüche als Ende eines Absatzes, was zur Folge hat, dass ganze Sätze beim Einlesen der Datei völlig willkürlich auseinandergerissen werden und eine weitere Analyse schlicht unmöglich ist.

Fallweise trifft man auf eine recht eigenwillige Markierung der direkten Rede, was ebenfalls zur Folge hat, dass manche Texte nicht von DCAnalyse verarbeitet werden können. Der Aufwand, diese Datei von Hand umzuformatieren, lohnt sich in der Regel nicht.

Auch in den E-Texten von Jules Verne, die vom *Project Gutenberg* stammen, treten merkwürdige Formatierungen auf, so z.B. in *L'Île mystérieuse* (1875). Der Text wird nicht nur nach ca. 70 Zeichen umgebrochen, sondern offenbar sind bei der Durchsicht des Dokuments nach Anwendung der OCR-Software nicht alle Sonderzeichen überprüft worden. Fett hervorgehobene Wörter wurden nicht konsequent in *Guillemets* gesetzt, sondern werden durch Bindestrich und sich schließende *Guillemets* markiert. Über die Entstehung dieser Fehlmarkierung lässt sich nur spekulieren, da sie aber an mehreren Stellen vorkommt, scheint es sich um einen systematischen Fehler zu handeln, z.B.:

... s'évasa et se rompit. Le vent, de l'état d'ouragan, passa au – **grand frais**», c'est-à-dire que la vitesse de translation des couches atmosphériques diminua de moitié. C'était encore ce que les marins appellent «une brise à trois ris», mais l'amélioration dans le trouble des éléments n'en fut pas moins considérable. C'était une baleine de la plus grande taille, qui, vraisemblablement, devait appartenir à l'espèce australe, dite – **baleine du Cap**». ⁵¹

Bei Texten, die als Hypertext Markup Language (HTML) vorliegen, ergeben sich ähnliche Ungereimtheiten: Was auf dem Bildschirm ansprechend formatiert aussieht – HTML bietet im Gegensatz zum reinen Textformat differenziertere Möglichkeiten zur Darstellung (z.B. Fett-/Kursivdruck, andere Schriftfarbe/-größe etc) – ist im Quelltext nicht zwangsläufig HTML-konform: So ist die Kursivdarstellung bei Gallica Classique in vielen Texten, v.a. aber bei Honoré de Balzac, fehlerhaft, indem der entsprechende Tag (<i>) zwar geöffnet, aber nicht wieder geschlossen wird (</i>). Da Web-Browser im allgemeinen bei der Anzeige von HTML-Code sehr fehlertolerant sind, merkt der Benutzer hiervon zwar nichts, jedoch können solche an sich unnötigen Ausnahmen für das Programm zu fehlerproduzierenden Hindernissen werden.

3.4.2 Durch Tagging induzierte Fehler bei der automatischen Auswertung

Eine weitere mögliche Fehlerquelle ist fehlerhafter Output von TreeTagger. Im Folgenden werden Beispiele für falsche Aufzeichnungen aufgeführt. Über die Entstehung der Falschauszeichnungen kann an dieser Stelle jedoch nicht genauer eingegangen werden.

Verschiedene Wortarten werden beim Taggen der redееinleitenden Sätze fälschlicherweise als Eigennamen identifiziert und mit dem Tag (NAM) versehen, z.B.:

(511) [Déjà (NAM)] [la (DET:ART)] [comtesse (NOM)] [Sabine (NAM)] [REPRENAIT (VER:impf)] [: (PUN)] ⁵²

(83) [Quand (NAM)] [ces (PRO:DEM)] [deux (NUM)] [vieillards (NOM)] [furent (VER:simp)] [dans (PRP)] [l' (DET:ART)] [embrasure (NOM)] [de (PRP)] [la (DET:ART)] [fenêtre (NOM)] [la (DET:ART)] [plus (ADV)] [éloignée (VER:pper)] [des (PRP:det)] [des (PRP:det)] [Grassins (NAM)] [: (PUN)] ⁵³

⁵¹ Jules Verne, *L'Île mystérieuse* (1875).

⁵² Émile Zola, *Nana* (1879).

⁵³ Honoré de Balzac, *Eugénie Grandet* (1843).

Da sich dies bei der später vorzunehmenden Gruppierung der einleitenden Sätze auswirkt, wo nach einem Adverb oder einer Konjunktion zu Beginn des Satzes gesucht wird, muss der Fehler vor der Suche korrigiert werden. Von dieser Falschauszeichnung betroffen sind außerdem die Adverbien *alors, aussi, aussitôt, cependant, déjà, enfin, ne, néanmoins, puis, sitôt*, die Konjunktionen *et, lorsque, mais, or, quand, si*, sowie verschiedene einzelne Präpositionen, Demonstrativa und Pronomina.

Ähnliches tritt beim Taggen der redebegleitenden Sätze auf. Dabei werden Eigennamen wie z.B. *Lélia, Lisbeth, Louisa, Luigi, Marneffe* als Verben und Eigennamen wie *Aline, Baudu, Dambreuse, Flore, Lorilleux, Michu, Roque, Surgère* als Adjektive identifiziert. In manchen *gérondif*-Konstruktionen wird außerdem das dem *participe présent* vorangehende *en* als Personalpronomen statt als Präposition ausgezeichnet, manchmal das *participe présent* auch als Adjektiv statt als *participe présent*. Innerhalb der Wendung *d'un ton* wird zuweilen *ton* als Possessivpronomen statt als Nomen ausgezeichnet.

Sehr häufig werden *Jamais, Je, Le, Ma, Mais, Oui* oder *Pas* am Anfang einer Zeile fälschlicherweise als Verb identifiziert. Ein Filter in DCAnalyse sorgt daher dafür, dass statt der Angabe VER:xxxx⁵⁴ die Tag-Bezeichnung TAGGER eingefügt wird, um anzuzeigen, dass dem Tagger an dieser Stelle ein Fehler unterlaufen ist (s. Bsp. (247)).

Ein weiterer Fehler unterläuft dem Tagger im folgenden Beispiel, indem der Eigennamen *Quenu* als *participe passé* identifiziert wird:

```
(247) [- (PUN)][ Mais (TAGGER)][, (PUN)][cria (VER:simp)][Quenu
(VER:pper)] [, (PUN)][jamais (ADV)][la (DET:ART)][charcuterie
(NOM)][du (PRP:det)] [vieux (NOM)][n' (ADV)][a
(VER:pres)][valu (VER:pper)][quinze (NUM)][mille
(NUM)][francs (NOM)][! (SENT)][Je (PRO:PER)][n' (ADV)][en
(PRO:PER)][aurais (VER:cond)][pas (ADV)][donné (VER:pper)]
[dix (NUM)][mille (NOM)][, (PUN)][moi (PRO:PER)][! (SENT)]55
```

Der Tagger scheint die Wortart in Abhängigkeit von der Position des Tokens zu bestimmen, denn ein paar Sätze zuvor wird *Quenu* korrekt ausgezeichnet:

```
(218) [- (PUN)][ Bonsoir (NOM)][, (PUN)][monsieur (NOM)][Quenu
(NAM)][. (SENT)]56
```

Die falsche Etikettierung bewirkt, dass DCAnalyse den Beispielsatz (247) nicht den redebegleitenden Sätzen zuordnet, sondern dieser fälschlicherweise zu den Sätzen einsortiert wird, die weder Redeeinleitung noch Redebegleitung enthalten.

⁵⁴ xxxx steht für die Angabe des Tempus.

⁵⁵ Émile Zola, *Le ventre de Paris* (1873).

⁵⁶ Ebenfalls Émile Zola, *Le ventre de Paris* (1873).

3.5 Programmende

Nachdem DCAnalyse alle Arbeitsschritte durchlaufen hat, ist die Analyse beendet, und das Programmfenster schließt sich. Die Exportdateien können nun im Unterverzeichnis `/output` eingesehen werden (siehe auch Kapitel 3.3.9).

Kapitel 4

Darstellung und Diskussion der Ergebnisse

Das Bearbeiten der im Korpus enthaltenen Texte mit dem Programm DCAnalyse, dessen Arbeitsablauf in Kapitel 3 beschrieben wurde, liefert umfangreiches Material sowohl zur direkten Rede als auch zum *discours citant*. Für die Auswertung der Daten und die Darstellung der Ergebnisse mittels Tabellen und Abbildungen wurden die Autoren stellenweise in Gruppen so zusammengefasst, dass die Publikationszeitpunkte ihrer Werke ungefähr beieinander liegen. Eine Ausnahme davon wird bei Victor Hugos Roman *Les misérables* (1862) gemacht, der um der Vergleichbarkeit willen in der Gruppe Stendhal, George Sand und Eugène belassen wurde, obwohl er zwölf Jahre später als der jüngste Text der Gruppe veröffentlicht wurde. Honoré de Balzac und Émile Zola bilden jeweils eine eigene Gruppe, da sie zu den sehr produktiven Autoren gehörten und die Datenlage für beide so gut ist, dass je 16 Romane in das Korpus aufgenommen werden konnten.

Innerhalb der Abbildungen 4.1 bis 4.4, 4.15 bis 4.31 und 4.32 bis 4.42 sind die Romantitel entlang der horizontalen Achse chronologisch sortiert und mit einer Nummer abgekürzt, während die Romane innerhalb der Graphiken 4.5 bis 4.14 über Kurztitel identifiziert werden. Die Kurztitel und Nummern sind anhand der Tabelle 3.1 auf Seite 45 über die vorletzte bzw. letzte Spalte aufzulösen. Folgende sechs Autorengruppen haben sich ergeben:

- Voltaire und Bernardin de Saint-Pierre
- Victor Hugo, Stendhal, Eugène Sue, George Sand
- Honoré de Balzac
- Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt, Gustave Flaubert
- Émile Zola
- Marcel Prévost, Henri Alain-Fournier, Marcel Proust, Colette

Im ersten Unterkapitel wird zunächst das Verhältnis der direkten Rede zum Erzähltext erläutert und die Verteilung von Redeeinleitung, Redebegleitung und alleinstehenden Turns innerhalb der direkten Rede besprochen. In einem weiteren Unterkapitel werden verschiedene Merkmale des redebegleitenden *discours citant* untersucht. Es wird ermittelt, welche *verba communicandi* bevorzugt von den Autoren eingesetzt werden. Anschließend wird dargestellt, wie häufig die bereits in Kapitel 2.2.3 vorgestellten Konstruktionstypen innerhalb der Redebegleitung vorkommen und welche Autoren welche *verba communicandi* darin bevorzugt einsetzen. Auch die Type-Token-Relation der *verba communicandi*, mittels derer sich feststellen lässt, bei welchen Autoren Redeverben besonders variantenreich zum Einsatz kommen, wird näher betrachtet. Außerdem wird die Frage erörtert, ob sich Verben, die innerhalb einer *gérondif*-Konstruktion das *verbum communicandi* semantisch erweitern, sich zu einem späteren Zeitpunkt selbst zu vollwertigen Redeverben entwickeln, wovon z.B. Vivienne Mylne¹ ausgeht. Das letzte Unterkapitel widmet sich den Arbeiten, die sich ebenfalls mit einzelnen zuvor genannten Aspekten beschäftigen und Zahlenwerte liefern. Diese werden besprochen und mit den im Zuge dieser Arbeit entstandenen Ergebnissen verglichen.

4.1 Allgemeine Eigenschaften der Figurenrede

Um einzelne Eigenschaften vor allem des redebegleitenden *discours citant* besser beurteilen zu können, ist es hilfreich, auch allgemeine Merkmale der ihm syntaktisch übergeordneten Figurenrede zu kennen. Dazu gehören vor allem das Verhältnis der direkten Rede zum Erzähltext sowie die Verteilung von Redeeinleitung, Redebegleitung und alleinstehenden Turns.

4.1.1 Verhältnis der direkten Rede zum Erzähltext

Bei der Berechnung des Anteils der direkten Rede am Text werden die für die direkte Rede ermittelten Tokens – dazu werden auch die Tokens innerhalb der Redeeinleitung und auch der Redebegleitung gezählt – von der Gesamttokenzahl des Textes subtrahiert. Redeeinleitung und Redebegleitung werden mit in die direkte Rede eingerechnet, weil die Anzahl der Tokens für die Redebegleitung, anders als für die Redeeinleitung, die durch den Doppelpunkt als graphischen Marker eindeutig identifizierbar ist, nicht zuverlässig bestimmt werden kann, was die beiden folgenden Beispielsätze demonstrieren:

¹ Vgl. Vivienne Mylne (1994), S. 42.

(13) – *Ah ! je sais, reprit Madeleine, c'est la femme qui est venue demeurer ici, qui est emménagée d'hier soir ?*²

(170) – *Voilà vingt francs, lui dit Fabrice, qui ne se sentait pas de joie de se trouver entre les jambes un cheval qui eût du mouvement.*³

Der *discours citant* ist in Beispiel (13) in die direkte Rede eingebettet (siehe Unterstreichung) und auf einen Teilsatz begrenzt. Nach dessen Ende, an dem ein Komma steht, wird direkte Rede im darauffolgenden Teilsatz fortgeführt. In Beispiel (170) liegt eine syntaktisch identische Konstruktion vor, jedoch gehören beide der direkten Rede folgenden Teilsätze zum *discours citant*, der sich in diesem Fall in Endstellung befindet. Bei dem ersten Teilsatz handelt es sich um den eigentlichen *discours citant*, der durch den zweiten Teilsatz um einen relativischen Nebensatz erweitert wird. Für ein Computerprogramm, das zwar Teilsätze und auch Wortarten bestimmen kann, ist es jedoch ohne Verständnis der Satzbedeutung unmöglich zu entscheiden, wo die direkte Rede tatsächlich endet bzw. an welcher Stelle diese nach einer Unterbrechung wiederaufgenommen wird.

Das Ergebnis der Subtraktion von Gesamttokenzahl und in direkter Rede ermittelter Tokens wird als Erzähltext angesehen.⁴ Abbildung 4.1 zeigt das Verhältnis von Tokens in direkter Rede zu Erzähltext-Tokens für alle im Arbeitskorpus enthaltenen Romane. Abbildung 4.2 bildet dieses Verhältnis um einer besseren Vergleichbarkeit willen noch einmal prozentual ab.

Anhand dieser beiden Abbildungen, die den Anteil der direkten Rede am Gesamttext illustrieren, lässt sich für den gesamten Untersuchungszeitraum eine Art Dreiteilung ausmachen. In den frühen Texten von Voltaire und Bernardin de Saint-Pierre spielt die direkte Rede noch keine allzu wichtige Rolle. Zwar ergeben sich für drei Romane von Voltaire, *Memnon* (1750), *Candide* (1759) und *Le blanc et le noir* (1764), Anteile zwischen 33,4% und 36,9%, jedoch liegen die übrigen Texte deutlich darunter. In *Paul et Virginie* (1787) beträgt der Anteil der Figurenrede sogar nur 8,2%, was innerhalb des Arbeitskorpus nur noch in *Salammbô* (1862) von Gustave Flaubert (8%) unterboten wird. Zwischen 1823 bis 1850 wird die direkte Rede dann deutlich häufiger eingesetzt als in den danach entstandenen Werken. Mit Ausnahme von drei Romanen (Nr. 11⁵, 12⁶ und 15⁷) beträgt

² George Sand, *François le Champi* (1850).

³ Stendhal, *La chartreuse de Parme* (1839).

⁴ Innerhalb des Erzähltextes kann wiederum indirekte und erlebte Rede enthalten sein, die jedoch in dieser Arbeit nicht als Untersuchungsgegenstand berücksichtigt werden können.

⁵ Victor Hugo, *Notre-Dame de Paris* (1831).

⁶ George Sand, *Indiana* (1832).

⁷ Honoré de Balzac, *La fille aux yeux d'or* (1835).

ihr Anteil mindestens 30%. In neun Romanen (Nr. 14⁸, 16⁹, 19¹⁰, 21¹¹, 26¹², 30¹³, 31¹⁴, 35¹⁵ und 37¹⁶) liegt er über 40%, in weiteren elf Romanen (Nr. 8¹⁷, 10¹⁸, 13¹⁹, 20²⁰, 22²¹, 23²², 25²³, 27²⁴, 29²⁵, 32²⁶ und 33²⁷) sogar bei 50% sowie deutlich darüber.

Auch wenn für die Romane Honoré de Balzacs aufgrund schwankender Werte kein eindeutiger Trend sichtbar wird, sind die Zahlen dennoch interessant: Für 15 von 16 analysierten Romanen macht der Anteil der direkten Rede mindestens 30% aus.²⁸ Insgesamt ergibt sich für Balzac ein Schnitt von 48%. Besonders redselig sind die Figuren in sechs Texten, die sich zu mehr als der Hälfte aus direkter Rede zusammensetzen: *La cousine Bette* (1848) mit 52,9%, *Le père Goriot* (1843) mit 55,6%, *Le colonel Chabert* (1844) mit 57,2%, *La peau de chagrin* (1831) mit 65,3% und *Le médecin de campagne* (1833) mit 77,5%.

In sechs weiteren Texten zwischen 1823 und 1850 lassen die Autoren ihre Figuren ebenfalls überdurchschnittlich viel reden. 50% und mehr beträgt der Anteil der direkten Rede in *Han d'Islande* (1823) von Victor Hugo, *Les mystères de Paris* (1842) und *Le juif errant* (1845) von Eugène Sue sowie *Les trois mousquetaires* (1844), *Le comte de Monte Christo* (1846) und *Le collier de la reine* (1849) von Alexandre Dumas. In drei von vier Romanen (Nr. 20²⁹, 27³⁰ und 29³¹), die mit einer Gesamtzahl von mehr als 500 000 Tokens sehr lang sind, ist deutlich mehr direkte Rede enthalten als Erzähltext. In Roman Nr. 40, *Les misérables* (1862) von Victor Hugo ist dieses Verhältnis jedoch mehr als

⁸ Honoré de Balzac, *La duchesse de Langeais* (1835).

⁹ Honoré de Balzac, *César Birotteau* (1837).

¹⁰ George Sand, *Consuelo* (1842).

¹¹ Honoré de Balzac, *Eugénie Grandet* (1843).

¹² Honoré de Balzac, *Le curé de village* (1845).

¹³ George Sand, *La mare au diable* (1846).

¹⁴ Honoré de Balzac, *Le cousin Pons* (1847).

¹⁵ Honoré de Balzac, *Les petits bourgeois* (1850).

¹⁶ George Sand, *François le Champi* (1850).

¹⁷ Victor Hugo, *Han d'Islande* (1823).

¹⁸ Honoré de Balzac, *La peau de chagrin* (1831).

¹⁹ Honoré de Balzac, *Le médecin de campagne* (1833).

²⁰ Eugène Sue, *Les mystères de Paris* (1842).

²¹ Honoré de Balzac, *Le père Goriot* (1843).

²² Honoré de Balzac, *Le colonel Chabert* (1844).

²³ Alexandre Dumas, *Les trois mousquetaires* (1844).

²⁴ Eugène Sue, *Le juif errant* (1845).

²⁵ Alexandre Dumas, *Le comte de Monte Christo* (1846).

²⁶ Honoré de Balzac, *La cousine Bette* (1848).

²⁷ Alexandre Dumas, *Le collier de la reine* (1849).

²⁸ *La fille aux yeux d'or* (1835) verfehlt diese Marke knapp mit 29,5%.

²⁹ Eugène Sue, *Les mystères de Paris* (1842): 64,1% direkte Rede.

³⁰ Eugène Sue, *Le juif errant* (1845): 53,7% direkte Rede.

³¹ Alexandre Dumas, *Le comte de Monte Christo* (1846): 59,5% direkte Rede.

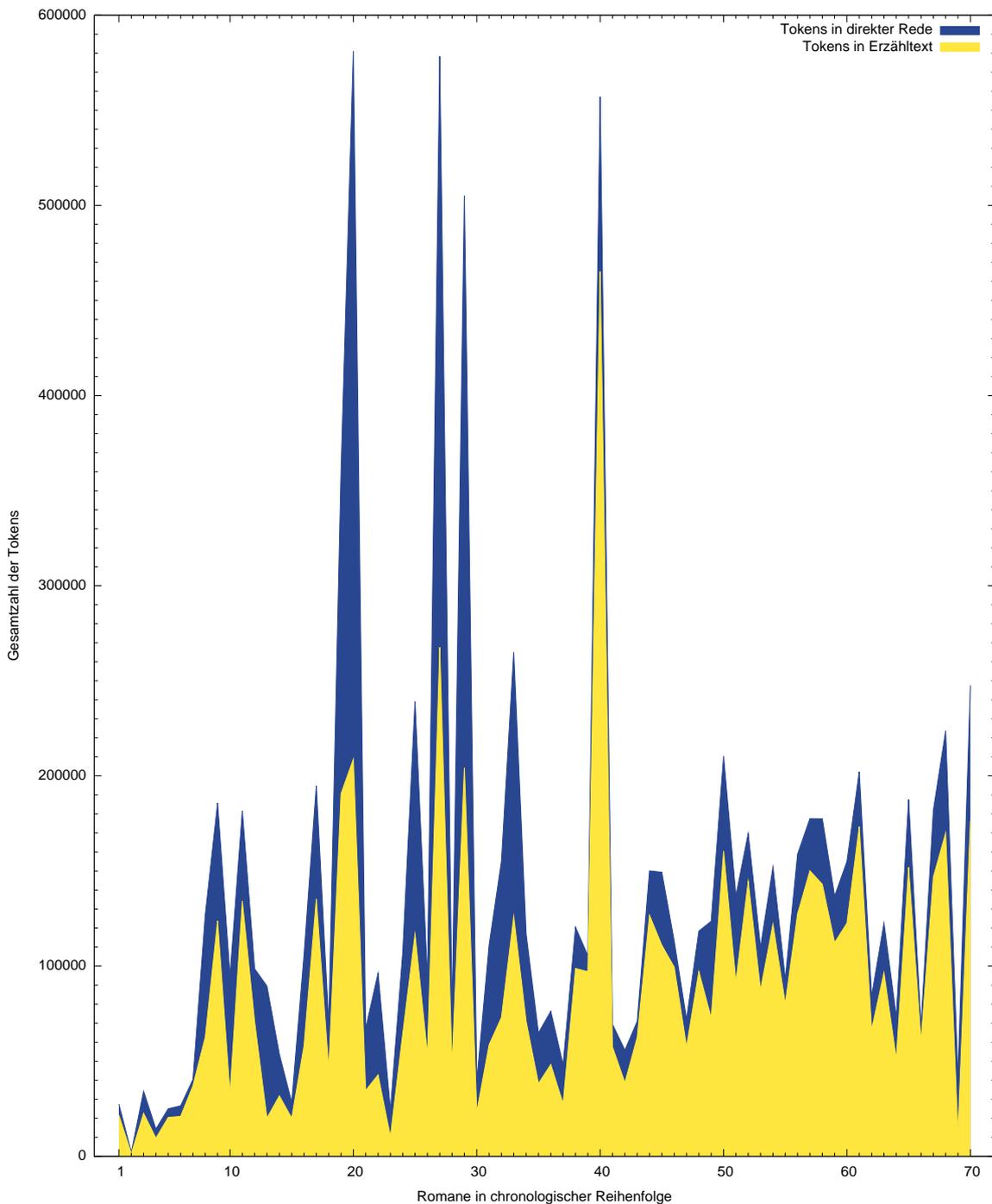


Abbildung 4.1: Verhältnis von direkter Rede und Erzähltext, gemessen in Tokens. Aus der ermittelten Gesamttokenzahl und der Anzahl der Tokens pro Turn lässt sich auf die restlichen Tokens, d.h. den Erzähltext, schließen. Die Werte für die gelbe und die blaue Fläche ergeben zusammen die Gesamtzahl der Tokens. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Um die Romane besser miteinander vergleichen zu können, wurde eine weitere Graphik erstellt, die den Anteil der direkten Rede am Text zusätzlich prozentual darstellt (s. Abbildung 4.2 sowie deren Erläuterung auf der nächsten Seite).

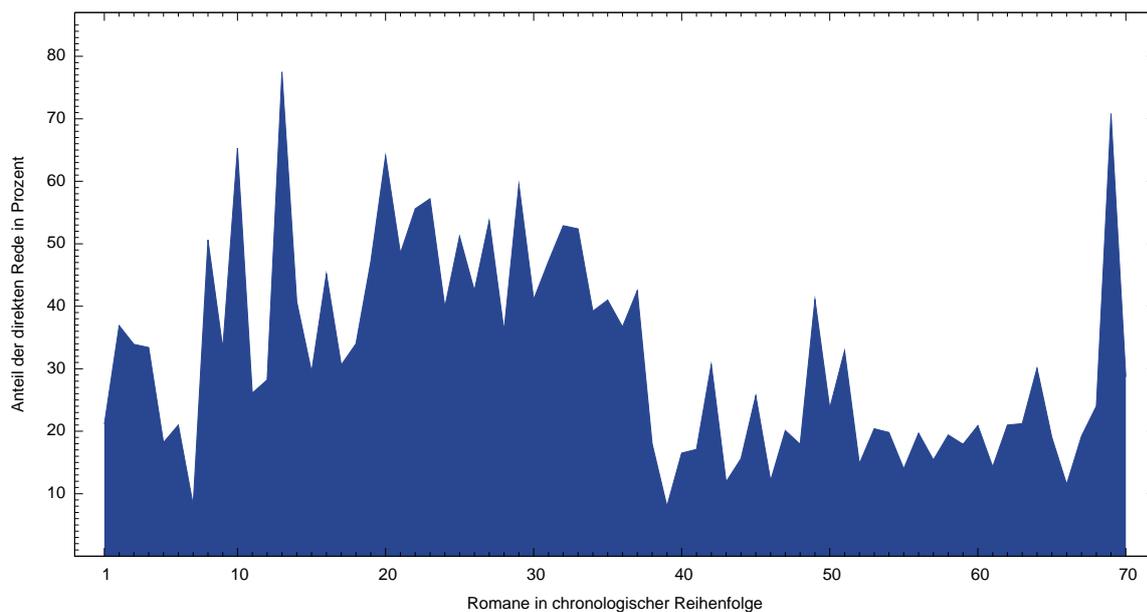


Abbildung 4.2: Anteil der direkten Rede am Erzähltext, gemessen in Prozent. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Für den gesamten Untersuchungszeitraum lässt sich eine Art Dreiteilung ausmachen. In den frühen Texten von Voltaire und Bernardin de Saint-Pierre spielt die direkte Rede noch keine allzu wichtige Rolle. Zwischen 1823 bis 1850 wird sie dann deutlich häufiger eingesetzt, so dass mit Ausnahme der Romane Nr. 11, 12 und 15 ihr Anteil mindestens 30% beträgt. Nach 1850 pendelt sich der Anteil der direkten Rede auf einem deutlich niedrigeren Niveau ein. Unter den Romanen Nr. 38-70 sind immerhin sechs (Nr. 42, 45, 49, 51, 64 und 69) vertreten, die 25% und mehr direkte Rede aufweisen. Der Anteil der Figurenrede innerhalb der restlichen 32 Romane ist mit Werten zwischen 8% und 23,7% hingegen deutlich niedriger und unterschreitet allein in 19 davon die 20%-Marke. Am Ende des Untersuchungszeitraums ragt allein *Chéri* (1920) von Colette mit einem ungewöhnlich hohen Anteil von 70,8% (Text Nr. 69) heraus.

umgekehrt: Von insgesamt 557 192 Tokens entfallen nur 16,5% auf die direkte Rede.

Nach 1850 pendelt sich der Anteil der direkten Rede auf einem deutlich niedrigeren Niveau ein. Unter den Romanen Nr. 38-70 sind immerhin sechs (Nr. 42³², 45³³, 49³⁴, 51³⁵, 64³⁶ und 69³⁷) vertreten, die 25% und mehr direkte Rede aufweisen. Der Anteil der Figurenrede innerhalb der restlichen 32 Romane ist mit Werten zwischen 8% und 23,7% hingegen deutlich niedriger und unterschreitet allein in 19 davon die 20%-Marke. Der Rückgang der direkten Rede ist wohl unter anderem auf die ab der zweiten Hälfte

³² Jules Verne, *De la terre à la lune* (1865): 30,7%.

³³ Jules Verne, *Ving-mille lieues sous les mers* (1869): 25,7%.

³⁴ Émile Zola, *La conquête de Plassans* (1874): 41,1%.

³⁵ Émile Zola, *Son Excellence Eugène Rougon* (1876): 32,9%.

³⁶ Marcel Prévost, *Les demi-vierges* (1894): 30,1%.

³⁷ Colette, *Chéri* (1920): 70,8%.

des 19. Jahrhunderts allmählich zunehmende Verwendung des *discours indirect libre* zurückzuführen.³⁸ Ausnahmen wie *Chéri* (1920) von Colette bestätigen jedoch die Regel.

4.1.2 Verteilung Redeeinleitung - Redebegleitung - alleinstehende Dialogstellen

Betrachtet man den *discours citant* gesondert nach Redeeinleitung, Redebegleitung und alleinstehenden Turns – deren Verteilung wird in den Abbildungen 4.3 und 4.4 graphisch dargestellt –, so lässt sich folgendes feststellen: Die Redeeinleitung spielt in der Mehrzahl der Korpustexte nur eine untergeordnete Rolle, da sie in 42 Romanen weniger als 10% beträgt. In nur 18 Texten übersteigt der Anteil 20%, in nur sieben Fällen 20%, und in nur drei Romanen geht jeder dritten Äußerung eine Redeeinleitung voraus³⁹. Der Anteil der Redeeinleitung ist innerhalb der ersten 37 Texte des Korpus besonders niedrig, da er mit Ausnahme von drei Romanen⁴⁰ stets weniger als 10% ausmacht. Innerhalb der zweiten Korpushälfte, d.h. für die nach 1850 entstandenen Romane, nimmt die Redeeinleitung allerdings leicht zu, was sich – sei es ein Zufall oder auch nicht – weitgehend mit der Beobachtung aus dem vorhergehenden Kapitel deckt, dass ab diesem Zeitpunkt die direkte Rede seltener eingesetzt wird. In den meisten der nach 1850 erschienenen Romane beträgt der Anteil der Redeeinleitung gerade dann 10% oder mehr, wenn der Anteil der direkten Rede gering ist und 20% nicht übersteigt.⁴¹

Ganz anders verhält es sich mit der Redebegleitung. Diese dominiert in 42 Texten mit einem Anteil von mindestens 50%, und auch in weiteren 14 liegt dieser über 40%. Die direkte Rede der restlichen 14 Texte steht jeweils zu mehr als 20% bzw. 30% in

³⁸ Vgl. Laurence Rosier (2008), S. 2.

³⁹ Siehe Bernardin de Saint-Pierre, *Paul et Virginie* (1787): 33%, Marcel Proust, *Du côté de chez Swann* (1913): 37% und *À l'ombre des jeunes filles en fleurs* (1918): 38%.

⁴⁰ Bernardin de Saint-Pierre, *Paul et Virginie* (1787): 33%, Stendhal, *La chartreuse de Parme* (1839): 14% und Honoré de Balzac, *Le lys dans la vallée* (1844): 14%.

⁴¹ Beispiele dafür (RE=Redeeinleitung, DD=<discours direct> / direkte Rede): Bernardin de Saint-Pierre, *Paul et Virginie* (1787): RE 33% / DD 8,2 %. Gustave Flaubert, *Madame Bovary* (1857): RE 12% / DD 18%. Gustave Flaubert, *Salammô* (1862): RE 24% / DD 8%. Victor Hugo, *Les misérables* (1862): RE 14% / DD 16,5%. Jules et Édmond de Goncourt, *Germinie Lacerteux* (1865): RE 20% / DD 17,1%. Émile Zola, *Thérèse Raquin* (1867): 16% / DD 11,9%. Gustave Flaubert, *L'éducation sentimentale* (1869): RE 13% / DD 15,6%. Émile Zola, *La curée* (1871): RE 25% / DD 12,1%. Émile Zola, *Le ventre de Paris* (1873): 26% / DD 17,9%. Émile Zola, *L'assommoir* (1877): RE 19% / DD 14,8%. Émile Zola, *Nana* (1879): RE 17% / DD 19,8%. Émile Zola, *Au bonheur des dames* (1883): RE 14% / DD 19,7%. Émile Zola, *Germinal* (1885): RE 11% / DD 15,4%. Émile Zola, *La bête humaine* (1890): RE 10% / DD 17,9%. Émile Zola, *La débâcle* (1892): RE 10% / DD 14,3%. Henri Alain-Fournier, *Le grand meaulnes* (1913): RE 15% / DD 11,5%. Marcel Proust, *Du côté de chez Swann* (1913): RE 37% / DD 19,2%.

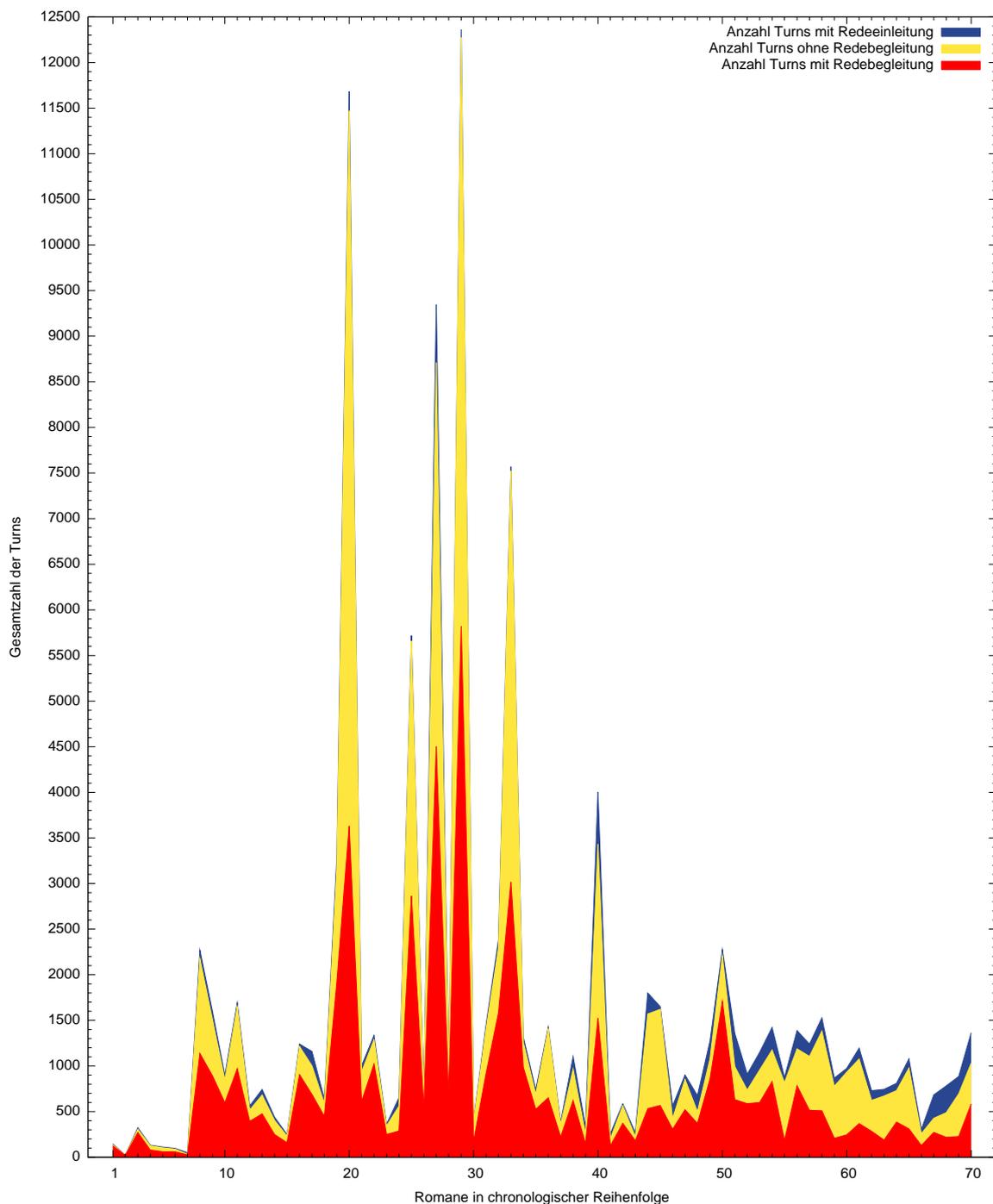


Abbildung 4.3: Verteilung von Redeeinleitung (blau), Redebegleitung (rot) und alleinstehenden Turns (d.h. ohne Redeeinleitung oder Redebegleitung, gelb) in allen Korpustexten. Die Werte der drei Farbflächen ergeben in der Summe jeweils die Gesamtzahl der pro Roman ermittelten Turns. Die Aufschlüsselung der Nummern für die Romane (x-Achse) erfolgt über Tabelle 3.1. Die sich ab der Korpusmitte andeutende Zunahme von Redeeinleitung und alleinstehenden Turns wird durch die prozentuale Darstellung der Verteilung in Abbildung 4.4 bestätigt.

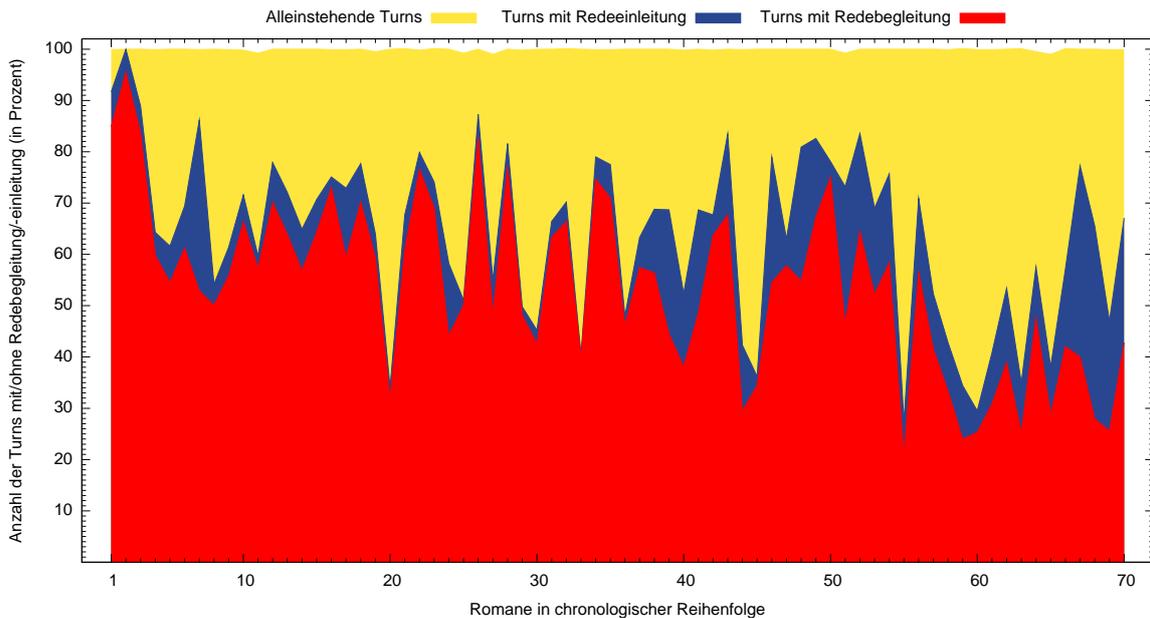


Abbildung 4.4: Verteilung von Redeeinleitung (blau), Redebegleitung (rot) und alleinstehenden Turns (d.h. ohne Redeeinleitung oder Redebegleitung, gelb) in allen Korpustexten. Die Werte der drei Farbflächen ergeben in der Summe jeweils 100% (die vereinzelten Abweichungen an der oberen x-Achse sind auf Rundungsfehler zurückzuführen). Die Aufschlüsselung der Romannummern entlang der x-Achse erfolgt über Tabelle 3.1. Die Redeeinleitung spielt in der Mehrzahl der Korpustexte nur eine untergeordnete Rolle, da sie in 42 Romanen weniger als 10% beträgt. Innerhalb der zweiten Korpushälfte, d.h. für die nach 1850 entstandenen Romane, nimmt sie leicht zu. Die Redebegleitung hingegen dominiert in 42 Texten mit einem Anteil von mindestens 50%. Bis 1879 liegt der Anteil der Redebegleitung in 54 Romanen bei mindestens 30%. Danach ist ein leichter Rückgang vor allem zugunsten alleinstehender Turns, aber auch der Redeeinleitung, zu beobachten. Alleinstehende Turns haben zwar insgesamt keinen so großen Anteil wie die Redebegleitung, sind aber sehr viel häufiger vertreten als die Redeeinleitung. In 14 Romanen steht jeder zweite Turn ohne Redeeinleitung oder Redebegleitung.

Verbindung mit einer Redebegleitung. Bis 1879 liegt der Anteil der Redebegleitung in 54 Romanen bei mindestens 30%. Danach ist ein leichter Rückgang vor allem zugunsten alleinstehender Turns, aber auch der Redeeinleitung, zu beobachten.

Alleinstehende Turns, denen also weder eine Redeeinleitung vorausgeht noch eine Redebegleitung folgt, haben zwar insgesamt keinen so großen Anteil wie die Redebegleitung, sind aber sehr viel häufiger vertreten als die Redeeinleitung. In 14 Romanen (Texte Nr. 20, 30, 33, 36, 44, 45, 55, 58, 59, 60, 61, 63, 65 und 69) steht mindestens jeder zweite Turn ohne Redeeinleitung oder Redebegleitung. Dass sich darunter allein sechs Romane von Émile Zola befinden, ist sicherlich auch weniger Zufall denn Stilmerkmal. Es kommt aber auch vor, dass sich in einzelnen Werken Redebegleitung und alleinstehende

Turns (innerhalb der Fußnoten mit RB und AT gekennzeichnet) nahezu die Waage halten (z.B. die Texte Nr. 8⁴², 24⁴³, 25⁴⁴, 27⁴⁵, 29⁴⁶ und 66⁴⁷). Die ausgeglichene Verteilung bewirkt jedoch auch, dass der Anteil der Redeeinleitung meist sehr gering ist und kaum mehr als 15% beträgt.⁴⁸ Besonders gegen Ende des Untersuchungszeitraums, d.h. in 13 von 16 Texten mit Entstehungsdatum zwischen 1881 und 1920 sind alleinstehende Turns stärker vertreten als redebegleitende Sätze. Das bedeutet, dass mit dem Rückgang der direkten Rede ab 1850 einerseits, der leichten Zunahme der Redeeinleitung sowie der Abnahme der Redebegleitung andererseits die alleinstehenden Turns populärer werden. Der Leser hat sich am Ende des 19. Jahrhunderts offenbar daran gewöhnt, dass nicht für jeden Turn die sprechende Figur explizit genannt werden muss.

Bei der Gegenüberstellung von Redeeinleitung, Redebegleitung und alleinstehenden Turns fallen außerdem die vier untersuchten Romane von Gustave Flaubert auf. In den beiden 1857 bzw. 1862 entstandenen Texten *Madame Bovary* und *Salammbô* überwiegt noch die Redebegleitung mit 56% und 44% im Vergleich zu 31% alleinstehenden Turns je Roman. Für die beiden späteren Romane aus den Jahren 1869 und 1881 hat sich das Verhältnis jedoch genau umgekehrt: In *L'éducation sentimentale* folgt der direkten Rede nur noch zu 30% eine Redebegleitung, während 72% aller Turns alleine stehen. In *Bouvard et Pécuchet* beträgt der Anteil der Redebegleitung sogar nur noch 21%, während die unbegleiteten Turns 72%, das Maximum innerhalb des Arbeitskorpus, ausmachen.

4.2 Merkmale des redebegleitenden *discours citant*

4.2.1 Verwendete *verba communicandi* in der Redebegleitung

Die Abbildungen 4.5 bis 4.13 stellen die Verteilung der innerhalb der Redebegleitung vorkommenden *verba communicandi*-Types (im Folgenden als VC-Types bezeichnet) dar. Für jeden Text jeder Autorengruppe wurden die zehn häufigsten *verba communicandi* ermittelt. Die Darstellung im Schaubild erfolgt zweifach: Die Abbildung links – es wurden

⁴² Victor Hugo, *Han d'Islande* (1823) (RB=Redebegleitung, AT=alleinstehende Turns): RB 50% / AT 46%.

⁴³ Honoré de Balzac, *Le lys dans la vallée* (1844): RB 44% / AT 42%.

⁴⁴ Alexandre Dumas, *Les trois mousquetaires* (1844): RB 50% / AT 48%.

⁴⁵ Eugène Sue, *Le juif errant* (1845): RB 48% / AT 45%.

⁴⁶ Alexandre Dumas, *Le comte de Monte Christo* (1846): RB 47% / AT 49%.

⁴⁷ Henri Alain-Fournier, *Le grand meaulnes* (1913): RB 42% / AT 43%.

⁴⁸ Beispiele: Victor Hugo, *Han d'Islande* (1823): RB 50% / AT 46%. Honoré de Balzac, *Le lys dans la vallée* (1844): RB 44% / AT 42%. Alexandre Dumas, *Les trois mousquetaires* (1844): RB 50% / AT 48%. Eugène Sue, *Le juif errant* (1845): RB 48% / AT 44%. Alexandre Dumas, *Le comte de Monte Christo* (1846): RB 47% / AT 49%. Henri Alain-Fournier, *Le grand meaulnes* (1913): RB 42% / AT 43%.

entweder zwei Grafiken auf eine Seite oder je eine Grafik auf eine einzelne Seite gedruckt – illustriert die absolute Verteilung, gemessen in Tokens, während die Darstellung rechts prozentual erfolgt. Auf diese Weise sind die Texte untereinander besser vergleichbar. Da nicht in allen Romanen dieselben zehn Kommunikationsverben am häufigsten sind, entsteht für jede Autorengruppe eine Vereinigungsmenge von Verben, die jeweils durch die Legende wiedergegeben wird. Um Platz für die prozentuale Darstellung zu gewinnen, wird die Legende für jede Autorengruppe nur jeweils einmal dargestellt: entweder innerhalb der Grafik mit zwei Schaubildern oder innerhalb der linken Grafik, sofern die Darstellung auf gegenüberliegenden Seiten erfolgt. Die Zahl in Klammern neben jedem Romankürzel gibt an, wieviele verschiedene VC-Typen insgesamt innerhalb der Redebegleitung ermittelt wurden. Die den Abbildungen zugrundeliegenden absoluten Zahlen sind den Tabellen A.1 bis A.6 in Anhang A zu entnehmen.

Voltaire und Bernardin de Saint-Pierre

Wie bereits Körner (1977), Vivienne Mylne (1994) und andere festgestellt haben, ist die Gruppe der Kommunikationsverben, die zur Ausgestaltung der Figurenrede eingesetzt werden, bis zum Ende des 18. Jahrhunderts überschaubar und geradezu monoton. Die Anzahl der von Voltaire und de Saint-Pierre verwendeten *verba communicandi* (VC) ist im Vergleich zu 50 Jahre später entstandenen Texten gering und schwankt zwischen sechs (in *Memnon* (1750) von Voltaire und *Paul et Virginie* (1787) von Bernardin de Saint-Pierre) und zwölf (*Zadig* (1747) von Voltaire) verschiedenen Typen.

Das bei Voltaire häufigste VC ist *dire*, gefolgt von *répondre*, *reprendre*, *s'écrier* und *répliquer*, von denen bis auf letzteres jeweils alle zu den häufigsten zehn gehören. Nicht ganz so häufig vertreten sind *continuer*, *crier*, *demander* und *ajouter*. Auffällig ist die Anzahl des Kommunikationsverbs *dire* in Voltaires *Candide* (1759). Die 214 *dire*-Tokens machen einen Anteil von 79,3% der VC-Tokens aus. Ähnlich frequent ist *dire* sonst nur in *Le blanc et le noir* (1764) mit 67,1% bzw. in *La princesse de Babylone* (1768) mit 63,3%. Auch *faire* kommt bei Voltaire als Kommunikationsverb vor, jedoch nur in dreien seiner Romane und dort auch nur sehr sparsam: In *Memnon* (1750), *Le blanc et le noir* (1764) und *L'ingénu* (1767) wurde nur je ein Token nachgewiesen.

In *Paul et Virginie* (1787) von Bernardin de Saint-Pierre hat die direkte Rede nicht nur einen sehr geringen Anteil von 8,2%, sondern jedes zweite VC ist außerdem ein *dire*-Token. Für minimale Abwechslung sorgen sonst nur *reprendre*, *répondre*, *s'écrier* und *demander*. Die letzten beiden kommen jeweils nur ein einziges Mal vor.

Beide Autoren verwenden fast ausschließlich semantisch neutrale Verben. Neben

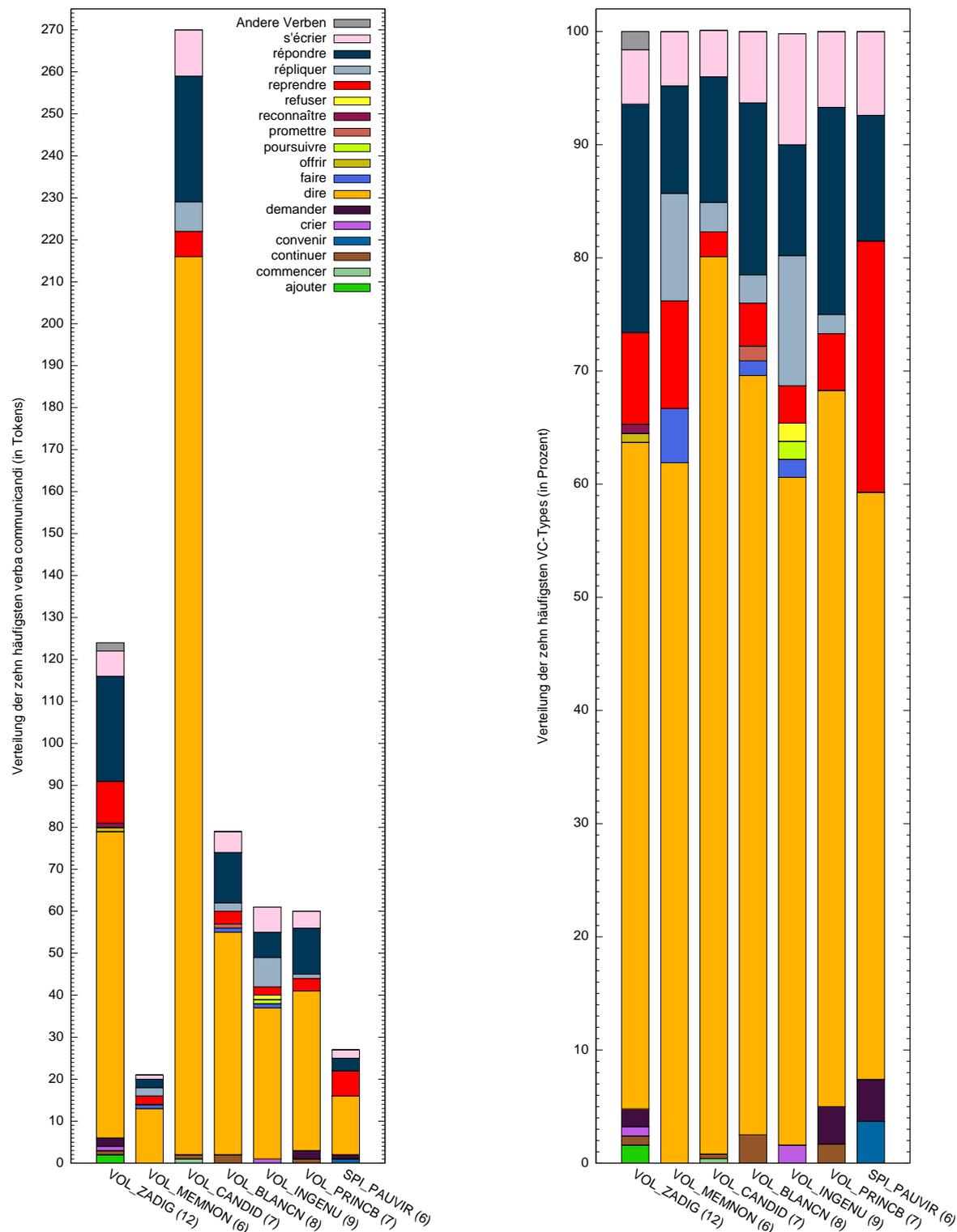


Abbildung 4.5: Die zehn häufigsten VC-Typen innerhalb der Redebegleitung bei **Voltaire** und **Bernardin de Saint-Pierre** in absoluter (links) und prozentualer (rechts) Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Roman-kürzel sind nach Abb. 3.1 aufzulösen.

dire werden wenige weitere Synonyma eingesetzt – die einzigen Verben jedoch, mit denen eine Modifikation der Stimme zum Ausdruck gebracht wird, sind *crier* und *s'écrier*.

Victor Hugo, George Sand, Stendhal und Eugène Sue

Mit dem Erscheinen von *Han d'Islande* von Victor Hugo im Jahre 1823 hat sich die Situation bezüglich der Kommunikationsverben grundlegend verändert. Liegt die Zahl der VC-Types innerhalb der Redebegleitung bei Voltaire noch zwischen sechs (*Memnon* (1750)) und zwölf (*Zadig* (1747)), ist diese bei Victor Hugo, George Sand, Stendhal und Eugène Sue bereits durchweg zweistellig, d.h. sie liegt zwischen zehn (George Sand, *La mare au diable* (1846)) und 85 (Eugène Sue, *Le juif errant* (1845)). Bei *Han d'Islande* (1823), *Notre-Dame de Paris* (1831) und *Les misérables* (1862) von Victor Hugo fällt auf, dass die Zahl der VC-Types für den über 30 Jahre später erschienenen Roman mit 47 Types deutlich größer ausfällt als für die beiden frühen Romane mit je 33 Types, obwohl in *Les misérables* die Anteile sowohl der direkten Rede (16,5%) als auch der Redebegleitung (38%) deutlich unter denen der beiden anderen liegen.⁴⁹

Zu den hochfrequenten, d.h. in allen Texten dieser Gruppe vorhandenen Kommunikationsverben gehören *dire*, *répondre*, *reprendre*, *s'écrier*, *faire* und *penser*. Der Anteil von *dire* bewegt sich zwischen 39,4% (*Han d'Islande* (1823) von Victor Hugo) und 59,9% (*Le rouge et le noir* (1830) von Stendhal), weshalb *dire* in allen Romanen das mit Abstand häufigste VC ist. Die Verben *répondre*⁵⁰, *reprendre*⁵¹ und *s'écrier*⁵² sind weitere häufige VC-Types, während *faire*⁵³ und *penser*⁵⁴ mit einzelnen Ausnahmen seltener eingesetzt werden.

Das Verb *faire*, das zwar bereits im Altfranzösischen als Kommunikationsverb Verwendung sowohl in der Redeeinleitung als auch in der Redebegleitung fand⁵⁵, jedoch zu Zeiten Voltaires und de Saint-Pierres kaum benutzt wurde, scheint sich in dieser Gruppe allmählich wieder als solches zu etablieren. Eine Zunahme der Verwendung von *faire*

⁴⁹ *Han d'Islande* (1823): 50,6% direkte Rede und 50% Redebegleitung. *Notre-Dame de Paris* (1831): 26,1% direkte Rede und 57% Redebegleitung.

⁵⁰ Der Anteil von *répondre* beträgt in dieser Gruppe zwischen 5,4% (*Les mystères de Paris* (1842) von Eugène Sue) und 22,7% (*La mare au diable* (1846) von George Sand).

⁵¹ *Reprendre* liegt mit Werten zwischen 5,1% (*La chartreuse de Parme* (1839) von Stendhal) und 15,4% (*Le juif errant* (1845) von Eugène Sue) knapp hinter *répondre*.

⁵² Das Kommunikationsverb *s'écrier* erreicht Werte zwischen 4,5% (*La mare au diable* (1846) von George Sand) und 16,7% (*Les mystères de Paris* (1842) von Eugène Sue).

⁵³ Für *faire* wurden Anteile zwischen 0,3% (*Han d'Islande* (1823) von Victor Hugo und *Indiana* (1832) von George Sand) und 9,9% (*François le Champi* (1850) von George Sand) ermittelt.

⁵⁴ *Penser* erreicht Werte zwischen 0,3% (*Han d'Islande* (1823) von Victor Hugo und *Le juif errant* (1845) von Eugène Sue) und 5,9% (*Le rouge et le noir* (1830) von Stendhal).

⁵⁵ Siehe Beispiel auf Seite 34.

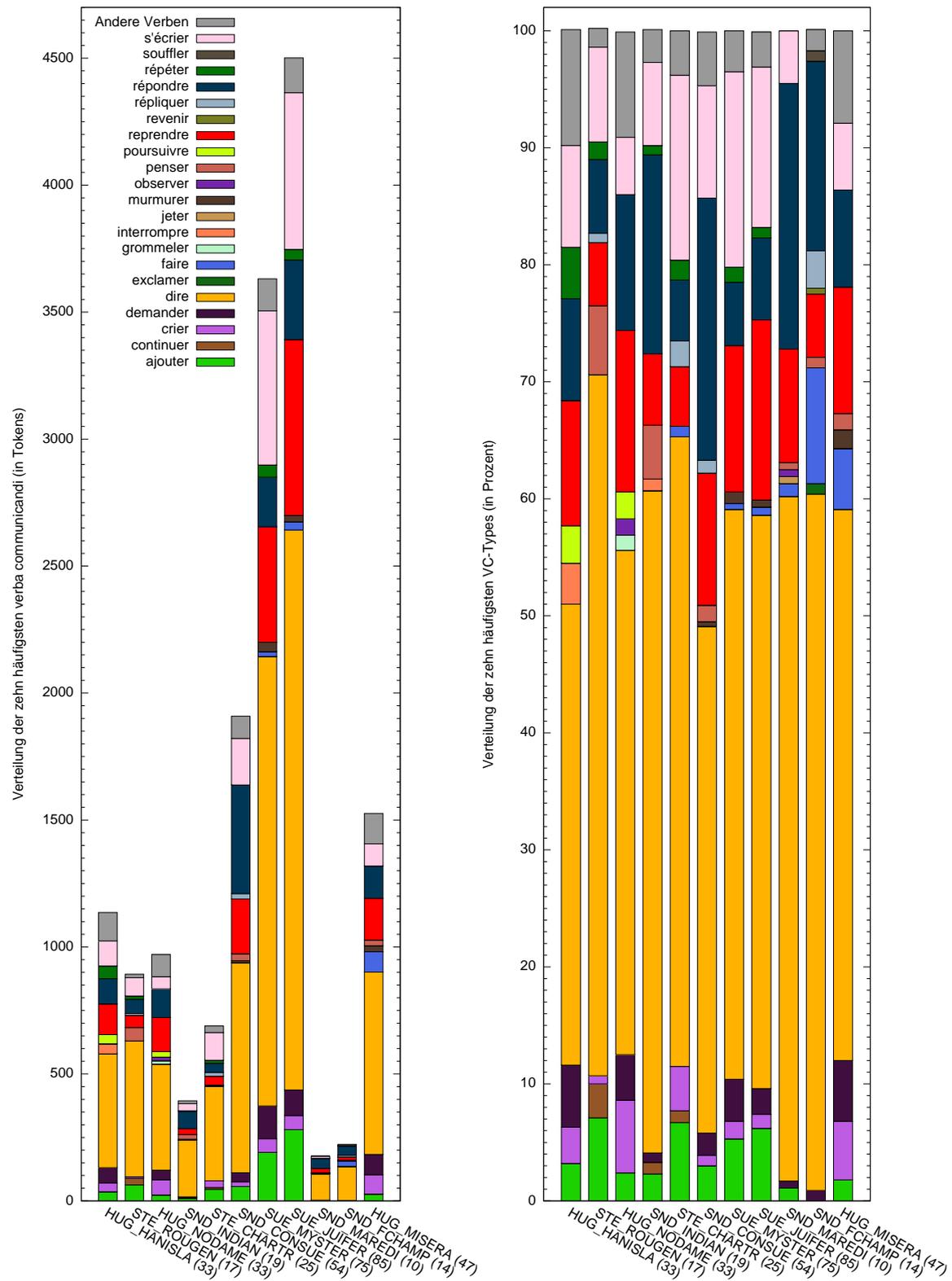


Abbildung 4.6: Die zehn häufigsten VC-Types innerhalb der Redebegeleitung bei **Victor Hugo**, **Stendhal**, **George Sand** und **Eugène Sue** in absoluter (links) und prozentualer (rechts) Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Types an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

als Kommunikationsverb ist vor allem bei Victor Hugo zu beobachten. In seinen beiden frühen Romanen *Han d'Islande* (1823) und *Notre-Dame de Paris* (1831) liegt der Anteil von *faire* nur bei 0,3% bzw. 0,4%, obwohl *Notre-Dame de Paris* um knapp 55 000 Tokens länger ist als *Han d'Islande*. *Les misérables* (1862) ist zwar knapp dreimal so lang wie *Notre-Dame de Paris*, jedoch hat der Anteil von *faire* mit 5,2% beträchtlich zugenommen. Dieser Wert wird nur noch mit 9,9% in *Indiana* (1832) von George Sand übertroffen.

Obwohl *Han d'Islande* (1823) von Victor Hugo der früheste Roman innerhalb der Gruppe ist, weist er mit 39,4% die wenigsten *dire*-Tokens auf, weshalb neben *dire* auch andere VC-Types höhere Anteile erreichen. Auf die restlichen 23 VC-Types verteilen sich 112 VC-Tokens. Anders verhält es sich mit Romanen wie *Le rouge et le noir* (1830) von Stendhal, *La mare au diable* (1846) und *François le Champi* (1850) von George Sand, in denen *dire* mit einem Anteil zwischen 58,5% (*La mare au diable*) und 59,9% (*Le rouge et le noir*) die restlichen neun häufigsten Verben nur eine marginale Rolle spielen. Da diese drei Romane zusätzlich nur zwischen zehn und 17 VC-Types insgesamt haben, entfallen auch nur jeweils wenige VC-Tokens – im Falle von *La mare au diable* sogar gar keine! – auf wenige Rest-VC-Types.⁵⁶

Eugène Sue konstruiert die Unterhaltungen seiner Figuren, indem er auffallend viele Kommunikationsverben einsetzt. Aufgrund der enormen Länge der beiden untersuchten Romane – *Les mystères de Paris* (1842) hat eine Gesamtlänge von 581 180 Tokens, *Le juif errant* (1845) 578 457 Tokens – und des hohen Anteils von direkter Rede (64,1% bzw. 53,7%) hat sich der Autor offensichtlich um eine abwechslungsreiche Redebegleitung bemüht. Die Anzahl von 75 bzw. 85 VC-Types ist um so bemerkenswerter, als vor allem in *Les mystères de Paris* nur jeder dritte Turn überhaupt eine Redebegleitung aufweist und sich die direkte Rede zu mehr als 65% aus alleinstehenden Turns zusammensetzt. Beide Romane haben außerdem dieselben zehn häufigsten VC-Types mit nur geringfügig voneinander abweichender Verteilung.

Der redebegleitende *discours citant* in den vier analysierten Romanen von George Sand ist mit Ausnahme von *Consuelo* (1842) vor allem durch eine geringe Gesamtlänge der Texte (zwischen 39 781 Tokens für *La mare au diable* (1846) und 95 587 Tokens für *Indiana* (1832)) und einer für diese Gruppe ungewöhnlich niedrigen Anzahl von VC-Types (zwischen zehn für *La mare au diable* (1846) und 19 für *Indiana* (1832)) gekennzeichnet. Anders als bei Eugène Sue, der in beiden untersuchten Texten fast dieselben Verben für die Redebegleitung einsetzt, folgen Sands Romane diesbezüglich keinem erkennbaren Schema. Zwar sind neben *dire* weitere drei Verben sehr häufig – darunter

⁵⁶ In *Indiana* (1832) verteilen sich 11 VC-Tokens auf neun VC-Types, in *François le Champi* (1850) sogar nur vier VC-Tokens auf vier VC-Types.

reprendre, *répondre* und *s'écrier* –, die restlichen sechs jedoch variieren von Text zu Text. So lässt Sand ihre Figuren in *Indiana* mit 4,6% überdurchschnittlich viel denken (*penser*)⁵⁷, während in *François le Champi* das Kommunikationsverb *faire* mit 9,9% den vierthöchsten Anteil innerhalb des Gesamtkorpus erreicht.

Honoré de Balzac

Das Arbeitskorpus enthält 16 Romane von Honoré de Balzac, deren Entstehungszeitraum von 1831 bis 1850 reicht. Die Anzahl der in den Texten ermittelten VC-Types bewegt sich zwischen zwölf (*La fille aux yeux d'or* (1835)) und 37 (*La peau de chagrin* (1831) und *Le médecin de campagne* (1833)). Dass mit zunehmender Textlänge auch die Anzahl der VC-Types innerhalb der Redebegleitung steigt, stimmt für Balzac nur bedingt: Während *La fille aux yeux d'or* mit 28 451 Gesamttokens mit nur zwölf VC-Types auskommt, sind es in dem mit 155 136 Tokens langen Roman *La cousine Bette* (1848) immerhin 23 VC-Types. Dennoch haben die beiden eingangs genannten Romane, *La peau de chagrin* und *Le médecin de campagne*, diesbezüglich ein deutlich besseres Verhältnis: Trotz einer geringeren Romanlänge von 94 940 bzw. 89 493 sind jeweils 37 VC-Types in der Redebegleitung gezählt worden. Das lässt sich dadurch erklären, dass beide Romane mit 65,5% und 77,5% einen sehr hohen Anteil von direkter Rede haben, die sich obendrein zu 66% bzw. 64% aus redebegleitenden Sätzen zusammensetzt.⁵⁸ Angesichts dieser Dialoglastigkeit scheint es nur die logische Konsequenz zu sein, dass sich der Autor mittels eines größeren Vorrats an Kommunikationsverben um Abwechslung bemüht.

Zu denjenigen VC, die in allen Romanen in großer Anzahl ermittelt wurden, gehören *dire*, *reprendre*, *répondre*, *s'écrier* und *demander*. Weit weniger frequent sind *ajouter*, *crier*, *répliquer*, *répéter* und *crier*. Balzac zieht außerdem *s'écrier* der Variante *crier* vor.

Das mit Abstand häufigste Kommunikationsverb ist *dire*, das hauptsächlich durch vier weitere Verben ergänzt wird: *répondre*, *reprendre*, *s'écrier* und *demander*. Die weiteren fünf der zehn häufigsten VC-Types werden allerdings nur noch durch wenige Tokens repräsentiert. Es fällt auf, dass der am frühesten veröffentlichte Roman, *La peau de chagrin* (1831), das ausgewogenste Verhältnis der VC-Types aufweist. Zwar ist *dire* auch hier das am häufigsten eingesetzte Kommunikationsverb, doch mit 33,6% liegt es deutlich unter dem Anteil in allen anderen Romanen. Zweithäufigstes *verbum communicandi*

⁵⁷ In dieser Gruppe hat mit 5,9% nur *Le rouge et le noir* (1830) von Stendhal noch mehr *penser*-Tokens.

⁵⁸ Zum Vergleich: *La fille aux yeux d'or* (1835) hat zwar einen ebenso großen Anteil von redebegleitenden Turns, doch beläuft sich die direkte Rede insgesamt nur auf 29,5%.

ist *s'écrier* mit 16,6%. Danach folgen *reprendre* und *répondre*. Zusammen mit *Le médecin de campagne* (1833) hat *La peau de chagrin* außerdem die meisten VC-Types (je 37), so dass sich 34 Tokens auf 27 weitere Types verteilen, die somit nur schwachfrequent sind. Anders verhält es sich für *César Birotteau* (1837). Obwohl in diesem Text mit 73% mehr Redebegleitung als in den zwei zuvor genannten enthalten ist (66% bzw. 64%), entfallen auf die Redebegleitung nur insgesamt 20 VC-Types, von denen *dire* mit 75,9% klar dominiert. Die restlichen neun häufigsten Kommunikationsverben verteilen sich auf nur 188 VC-Tokens. Darüber hinaus sind nur noch 10 weitere VC-Types enthalten, auf die jeweils ein VC-Token entfällt.

Das Verb *faire* ist als Kommunikationsverb zwar in allen Romanen enthalten, jedoch ist es zum einen nicht in allen Romanen unter den häufigsten zehn Verben und spielt zum anderen mit Werten zwischen 0,4% (*Le colonel Chabert* (1844)) und 2,7% (*Le médecin de campagne* (1833)) eine nur untergeordnete Rolle:

(718) – *Kt, kt, kt, kt, fit Bianchon en faisant claquer sa langue contre son palais, comme pour exciter un cheval.*⁵⁹

Balzacs lässt seine Figuren, anders als andere Autoren, in vielen seiner Romane per direkter Rede mitteilen, was sie denken. In neun Romanen gehört *penser* zwar zu den zehn häufigsten Kommunikationsverben, was aber nicht darüber hinwegtäuschen darf, dass die Anteile insgesamt mit zwischen 0,4% (*La femme de trente ans* (1842), *Le père Goriot* (1843), *Le colonel Chabert* (1844) und *La cousine Bette* (1848)) und 1,6% (*Eugénie Grandet* (1843)) gering sind.

(1131) – *Tu bats la breloque, mon vieux, pensa du Tillet en se servant d'un mot de son premier métier.*⁶⁰

(715) – *Peste ! je ne suis plus monsieur de Bonfons, pensa tristement le magistrat dont la figure prit l'expression de celle d'un juge ennuyé par une plaidoirie.*⁶¹

Diejenigen Kommunikationsverben, die über die zehn häufigsten VC-Types hinaus verwendet werden, sind je nach Roman mit zwischen zwei (*La fille aux yeux d'or* (1835) und 27 Types (*La peau de chagrin* (1831) und *Le médecin de campagne* (1833)) unterschiedlich zahlreich. Jedoch gilt, unabhängig von der jeweiligen Anzahl der Types, dass auf die VC-Types entweder nur ebenso viele oder geringfügig mehr VC-Tokens entfallen. Damit ist ausgeschlossen, dass eines dieser Verben auch nur annähernd so frequent ist wie die zehn häufigsten.

⁵⁹ Honoré de Balzac, *Le père Goriot* (1843).

⁶⁰ Honoré de Balzac, *César Birotteau* (1837).

⁶¹ Honoré de Balzac, *Eugénie Grandet* (1843).

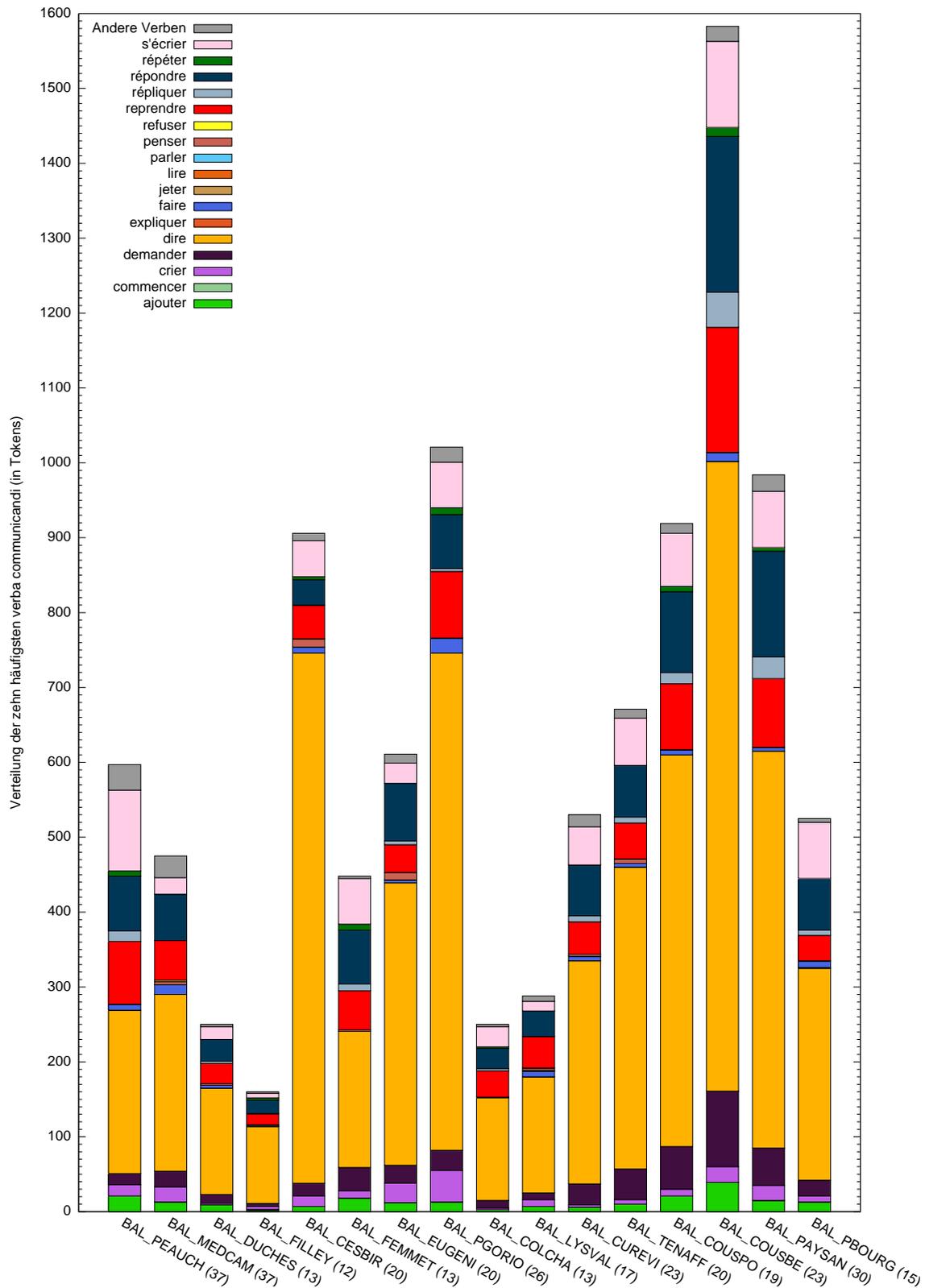


Abbildung 4.7: Die zehn häufigsten VC-Typen innerhalb der Redebegleitung bei **Honoré de Balzac** in absoluter Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

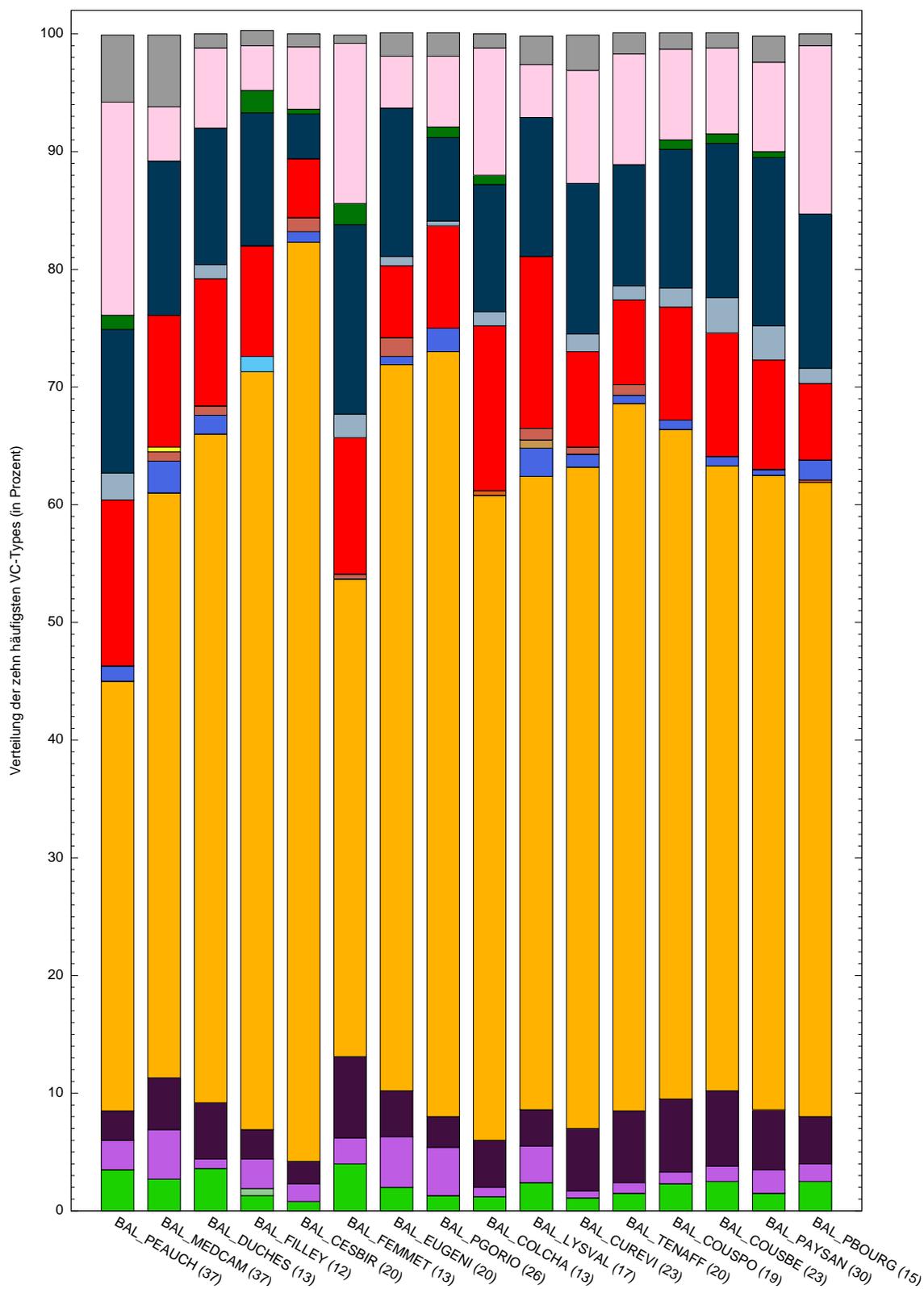


Abbildung 4.8: Die zehn häufigsten VC-Typen innerhalb der Redebegeleitung bei **Honoré de Balzac** in prozentualer Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt, Gustave Flaubert

In der Gruppe Dumas-Verne-de Goncourt-Flaubert kommen zwischen 15 (Jules Verne, *Le tour du monde* (1873)) und 67 (Alexandre Dumas, *Le comte de Monte Christo* (1846)) verschiedene VC-Typen vor. Von diesen sind acht Verben in allen Texten enthalten: *dire*, *répondre*, *s'écrier*, *repandre*, *faire*, *murmurer*, *répéter* und *crier*. Die drei Verben *demander*, *répliquer* und *ajouter* wurden bis auf eine Ausnahme (*Germinie Lacerteux* (1865)) ebenfalls in allen Romanen ermittelt. Weitere frequente VC, die jedoch nicht in allen Texten vorkommen, sind *continuer*, *penser*, *balbutier*, *poursuivre* und *interrompre*.

Während bei Flaubert, Dumas und den Gebrüdern de Goncourt *dire* in mindestens 40% (Gustave Flaubert, *Madame Bovary* (1857)) der Fälle – das Maximum liegt bei 55,1% (Alexandre Dumas, *Le comte de Monte Christo* (1846)) – das dominierende VC ist, liegt sein Anteil in den vier Texten von Jules Verne nur zwischen 17,4% (*Vingt-mille lieues sous les mers* (1869)) und 24,5% (*L'île mystérieuse* (1875)). Jules Verne ist damit der erste Autor innerhalb des Arbeitskorpus, der *dire* innerhalb der Redebegleitung nicht den Vorrang lässt. Stattdessen wird *dire* von *répondre* mit Anteilen von mindestens 28,8% (*De la terre à la lune* (1865)) und höchstens 39,5% (*L'île mystérieuse* (1875)) von der Spitzenposition verdrängt. Die mögliche Erwartung, dass bei häufiger Verwendung des Verbs *répondre* ein komplementäres Kommunikationsverb wie *demander* in ebenso großer Zahl vorhanden sein könnte, erfüllt sich jedoch nicht: *demander* kommt bei Jules Verne zwar auf einen Anteil von 7,8% (*De la terre à la lune* (1865)) bis 13,6% (*Vingt-mille lieues sous les mers* (1869)), doch weder *demander* selbst noch weitere dazu synonyme *verba communicandi* kommen in entsprechender Zahl vor.

Alexandre Dumas' vier untersuchte Romane zeichnen sich in drei von vier Fällen vor allem durch eine enorme Textlänge und eine dementsprechende Vielzahl von VC-Typen aus. Dennoch vergrößert sich die Anzahl der VC-Typen nicht proportional zur Textlänge. Während *Le comte de Monte Christo* (1846) mit 505 014 Gesamttokens der längste Text innerhalb dieser Gruppe ist und in der Redebegleitung 67 VC-Typen vorkommen, ist *Le collier de la reine* (1849) mit 265 042 Gesamttokens nur fast halb so lang und enthält dennoch 63 VC-Typen. *La tulipe noire* (1850) ist mit 76 269 Gesamttokens der kürzeste dieser vier Romane, weist aber, obwohl der Text um zwei Drittel kürzer ist als *Les trois mousquetaires* (1844)⁶² mit 31 statt 41 nur unwesentlich weniger VC-Typen auf, was um so bemerkenswerter ist, als der Anteil der direkten Rede mit nur 36,7% rund 15% unter dem des längeren Textes liegt. Während *Les trois mousquetaires* nur um knapp

⁶² *Les trois mousquetaires* (1844) von Alexandre Dumas enthält 239 109 Gesamttokens.

26 000 Tokens kürzer ist als *Le collier de la reine* und der Anteil der direkten Rede ungefähr gleich ist (51,5% zu 52,4%), der Anteil der Redebegleitung mit 50% jedoch um 10% höher liegt, ist dennoch die Zahl der verwendeten Kommunikationsverben mit 41 gering. Ein ähnlich ungewöhnliches Verhältnis liegt auch bei Jules Verne vor. Der Roman *L'île mystérieuse* (1875) ist mit 210 369 Gesamttokens zwar um ein gutes Drittel länger als *Vingt-mille lieues sous les mers* (1869), hat aber bei einem vergleichbaren Anteil von direkter Rede (23,7% zu 25,7%) und einem mehr als doppelt so großen Anteil von redebegleitenden Sätzen (75% zu 34%) mit insgesamt 22 VC-Types sogar vier weniger als der kürzere Text.

Innerhalb dieser Autorengruppe fällt *faire* erstmalig als Kommunikationsverb ins Auge: Es ist einerseits in allen Romanen enthalten und gehört andererseits in allen Texten zu den zehn häufigsten VC-Types. Die Anteile liegen zwischen 0,9% (Jules Verne, *L'île mystérieuse* (1875)) und 23,8% (Jules und Édmond de Goncourt, *Germinie Lacerteux* (1865)). Die Werte bei Jules Verne bewegen sich zwischen 0,9% (*L'île mystérieuse* (1875)) und 5,6% (*Vingt-mille lieues sous les mers* (1869)) und zeigen damit, dass der Autor *faire* sicherlich nicht als adäquaten Ersatz für *dire* angesehen hat. Auf die Gebrüder de Goncourt scheint dies schon eher zuzutreffen, da *dire* zwar Teil jeder zweiten Redebegleitung ist, aber *faire* mit 23,8% immerhin den größten Anteil innerhalb des Korpus ausmacht und damit für diesen Roman das zweithäufigste Kommunikationsverb darstellt. Während *faire* auch in den beiden frühen Romanen von Flaubert, *Madame Bovary* (1857) und *Salambô* (1862), zahlreich vertreten ist, geht der Anteil für *L'éducation sentimentale* (1869) und erst recht für *Bouvard et Pécuchet* (1881) mit 2,8% bzw. 1,1% wieder deutlich zurück.

Alexandre Dumas setzt im Vergleich zu den anderen Autoren dieser Gruppe bevorzugt die Verben *continuer* und *murmurer* ein. So erreicht *continuer* Werte zwischen 3,4% (*Le comte de Monte Christo* (1846)) und 5,2% (*La tulipe noire* (1850)), während das Verb sonst nur bei Gustave Flaubert und dort mit Anteilen von 1,9% und darunter viel seltener nachgewiesen wurde. Obwohl *murmurer* in allen Romanen dieser Gruppe vorkommt und sich zusätzlich bei Gustave Flaubert und Jules Verne unter den zehn häufigsten Kommunikationsverben befindet, überschreitet es nur bei Alexandre Dumas die Zwei-Prozent-Marke und erreicht in *La tulipe noire* (1850) sogar einen Anteil von 4%.

Gustave Flaubert gewährt vor allem in seinen beiden Romanen *Madame Bovary* (1857) und *L'éducation sentimentale* (1869) Einblicke in die Gedankenwelt seiner Figuren, indem er *penser* als Kommunikationsverb einsetzt. Mit 2,3% bzw. 2,1% ist es Teil der zehn häufigsten Verben innerhalb der Redebegleitung, wie auch in *Le tour du monde en*

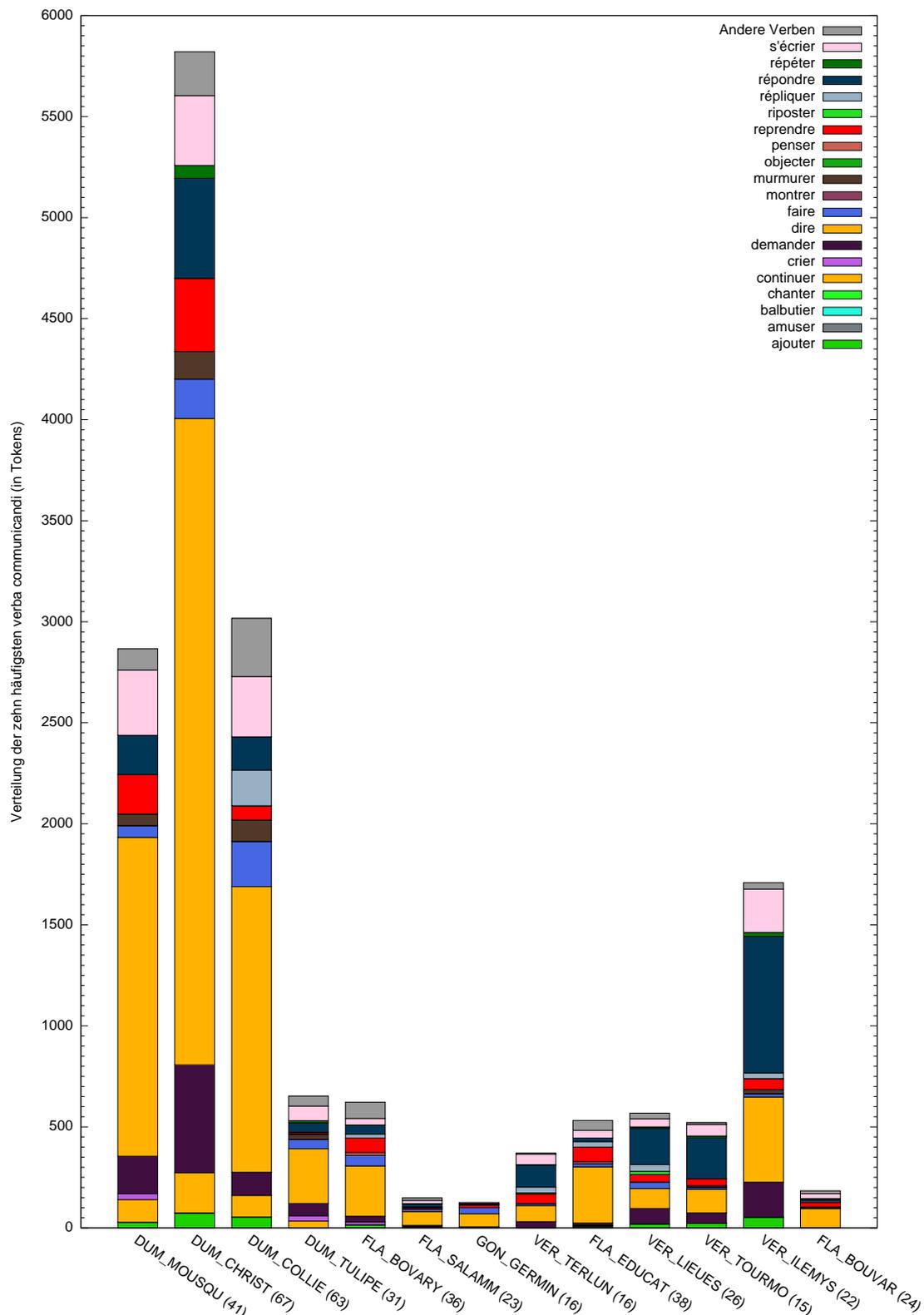


Abbildung 4.9: Die zehn häufigsten VC-Typen innerhalb der Redebegleitung bei **Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt** und **Gustave Flaubert** in prozentualer Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

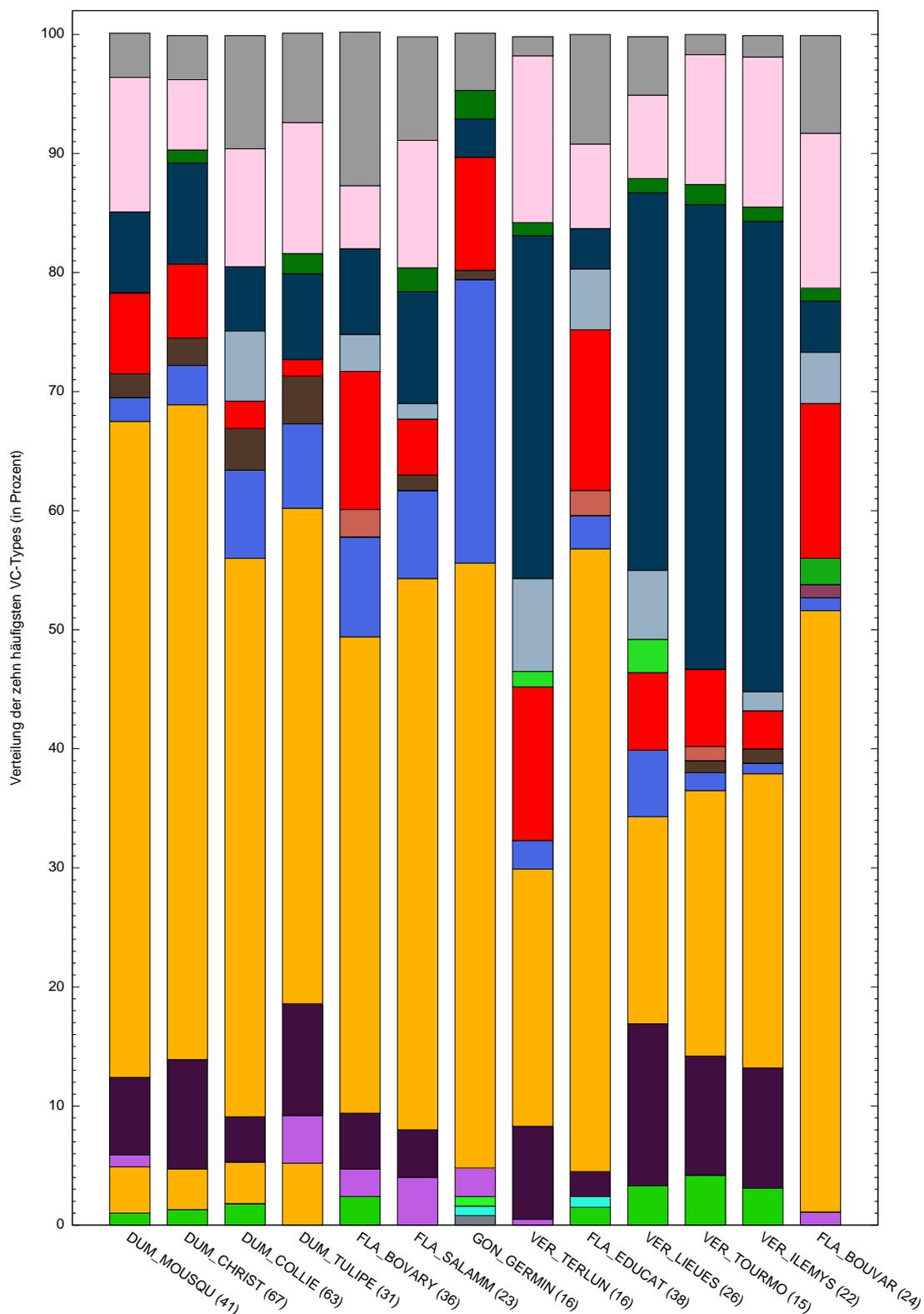


Abbildung 4.10: Die zehn häufigsten VC-Types innerhalb der Redebegleitung bei **Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt** und **Gustave Flaubert** in prozentualer Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Types an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

quatre-vingts jours (1873) von Jules Verne, wo es jedoch mit 1,2% deutlich seltener vorkommt. In den beiden anderen untersuchten Romanen von Flaubert, *Salammô* (1862) und *Bouvard et Pécuchet* (1881), ist es hingegen gar nicht enthalten.

Auch in dieser Autorengruppe dominiert *s'écrier* gegenüber *crier*. Allerdings fällt auf, dass *crier* noch seltener zu den zehn häufigsten Kommunikationsverben gehört als beispielsweise bei Honoré de Balzac oder der Gruppe Victor Hugo, George Sand, Stendhal und Eugène Sue. Bei Jules Verne taucht es in dieser Liste in drei von vier untersuchten Romanen gar nicht auf und ist mit Werten zwischen 0,4% und 0,5% insgesamt nur schwach vertreten. Das Maximum für *crier* beträgt 4% und wurde in *La tulipe noire* (1850) von Alexandre Dumas und in *Salammô* (1862) von Gustave Flaubert ermittelt. Das *verbum communicandi* *s'écrier* gehört mit Anteilen zwischen 5,9% (Alexandre Dumas, *Le comte de Monte Christo* (1846)) und 13% (Gustave Flaubert, *Bouvard et Pécuchet* (1881)) in fast allen Romanen dieser Gruppe zu den hochfrequenten Verben. Lediglich *Germinie Lacerteux* (1865) der Gebrüder de Goncourt fällt hier aus dem Rahmen, wo *s'écrier* mit lediglich 0,8% vertreten ist – das entspricht einem einzigen *s'écrier*-Token.

Betrachtet man schließlich die Kommunikationsverben, die über die häufigsten zehn VC-Types hinaus enthalten sind, so ist deren Anzahl mit zwischen fünf (Jules Verne, *Le tour du monde* (1873)) und 57 VC-Types (Alexandre Dumas, *Le comte de Monte Christo* (1846)) auch für diese Gruppe sehr unterschiedlich. Besonders in den Romanen, die insgesamt viele VC-Types aufweisen (z.B. bei Alexandre Dumas), entfallen auf die restlichen VC-Types viele VC-Tokens, weshalb es durchaus vorkommen kann, dass der an elfter Stelle liegende VC-Type fast ebenso frequent ist wie der zehnte.⁶³ Da aber in der Regel bereits der an zehnter Position liegende VC-Type schwachfrequent ist, ist auch für die nachfolgenden Types nicht mit Überraschungen hinsichtlich ihrer Häufigkeit zu rechnen.

Émile Zola

Die Anzahl der verschiedenen VC-Types in den 16 analysierten Romanen von Émile Zola reicht von 21 (*Thérèse Raquin* (1867)) bis zu 54 VC-Types (*Son Excellence Eugène Rougon* (1876)).

⁶³ Siehe z.B. *Le collier de la reine* (1849) von Alexandre Dumas: Jenseits der zehn häufigsten VC-Types sind weitere 53 Types enthalten, auf die sich insgesamt 288 Tokens verteilen. Der zehnthäufigste Type ist *ajouter* mit 54 Tokens. Für *répéter* wurden 53 Tokens ermittelt. Da *ajouter* mit einem Anteil von 1,8% nicht besonders häufig ist, liegen auch die Werte für alle nachfolgenden VC-Types entweder auf demselben Niveau oder darunter.

Die Gruppe der VC, die in allen Texten vorkommen, umfasst zehn Verben: *continuer, crier, demander, dire, déclarer, murmurer, reprendre, répondre, répéter* und *s'écrier*. Wie sich anhand der Vereinigungsmenge in Abbildung 4.11 feststellen lässt, handelt es sich dabei jedoch nicht um die zehn häufigsten eines jeden Romans. Weitere, in fast allen Romanen vorkommende Kommunikationsverben sind *ajouter, balbutier, bégayer, conclure, expliquer* und *interrompre*. Zwar variiert die Länge der 16 untersuchten Texte deutlich, jedoch fällt bei Betrachtung der Abbildung 4.12 auf, dass die Auswahl der in der Redebegleitung verwendeten Kommunikationsverben bei Émile Zola, anders als bei seinen Schriftstellerkollegen, sehr homogen ist, wenn man von einzelnen, nur kleineren Ausnahmen absieht. Zwar ist auch bei Zola *dire* das vorherrschende Kommunikationsverb, aber aufgrund von vergleichsweise moderaten Anteilen zwischen 17,9% (*L'argent* (1891)) und 38,8% (*La curée* (1871)), sind auch andere Verben stärker vertreten. Vom ältesten (*Thérèse Raquin* (1867): 32,2%) bis zum jüngsten untersuchten Roman (*Les trois villes* (1894): 19,4%), ist allerdings eine Abnahme des Anteils von *dire* zugunsten anderer Verben wie *reprendre, répéter, murmurer* sowie vereinzelt *déclarer* und *bégayer* zu beobachten, die erstmalig bei Zola zu den zehn häufigsten Kommunikationsverben gehören⁶⁴:

(1755) – *Non, merci, je n'ai besoin de rien, déclara Mme Marty.*⁶⁵

(1630) – *Quoi donc ? quoi donc ? bégaya Bazouge, qui est-ce qui se trouve mal ?
... On y va, la petite mère !*⁶⁶

Weitere, bisher in keiner anderen Autorengruppe zu den häufigsten zählende bzw. nur in einzelnen Texten nachgewiesene Verben sind *conclure*⁶⁷ und *balbutier*⁶⁸. Auch *hurler* ist dieser Reihe hinzuzufügen, jedoch gehört es nur in einem einzigen Roman, *Germinal* (1885), mit 1,9% zu den häufigsten VC-Types:

⁶⁴ In insgesamt sechs Romanen ist *déclarer* Teil der zehn häufigsten Kommunikationsverben, d.h. in *Son Excellence Eugène Rougon* (1876): 1,7%, *Nana* (1879): 2,2%, *Au bonheur des dames* (1883): 2,4%, *La terre* (1887): 5,9%, *L'argent* (1891): 5,7% und *La débâcle* (1892): 2,7%. *bégayer* gehört in zwei Texten zu den häufigsten Verben, und zwar in *L'assommoir* (1877): 2,2% und *Nana* (1879): 1,6%.

⁶⁵ Émile Zola, *Au bonheur des dames* (1883).

⁶⁶ Émile Zola, *L'assommoir* (1877).

⁶⁷ Émile Zola setzt *conclure* in seinen jüngeren Romanen *La terre* (1887), *La bête humaine* (1890), *Le docteur Pascal* (1893) und *Les trois villes* (1894) verstärkt ein, so dass das Verb mit Anteilen zwischen 1,9% und 2,8% zu den zehn häufigsten VC-Types gehört.

⁶⁸ Zuvor kommt *balbutier* nur in *Germinie Lacerteux* (1865) der Gebrüder de Goncourt und in *L'éducation sentimentale* (1869) von Gustave Flaubert und mit Anteilen von 0,8% bzw. 0,9% außerdem nur sehr selten vor. Bei Émile Zola ist *balbutier* vor allem in den älteren Texten enthalten und in fünf Romanen mit Anteilen zwischen 1,3% (*La curée* (1871) bzw. *La conquête de Plassans* (1874)) und 2,8% (*Thérèse Raquin* (1867)) deutlich häufiger anzutreffen.

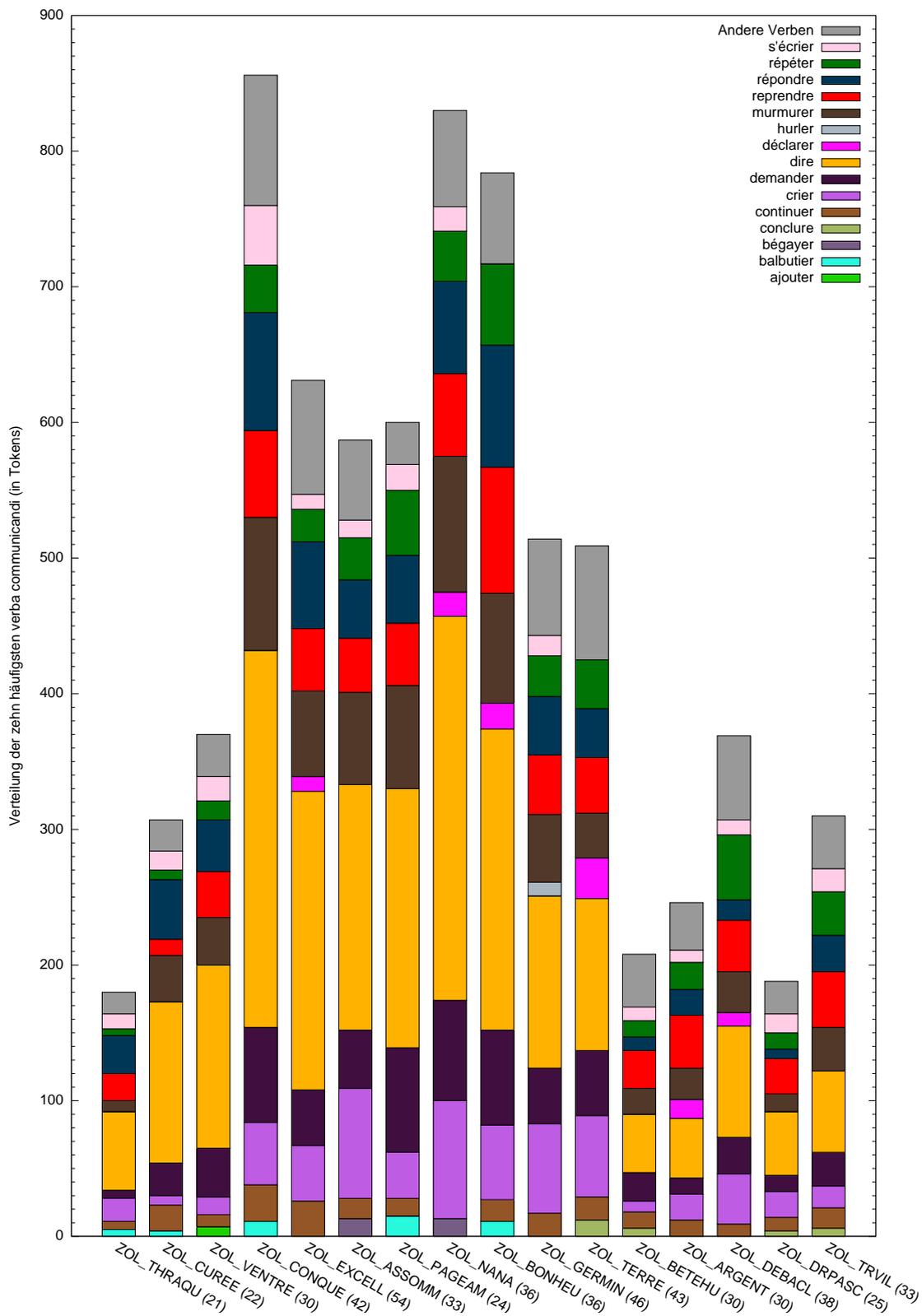


Abbildung 4.11: Die zehn häufigsten VC-Typen innerhalb der Redebegleitung bei **Émile Zola** in absoluter Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

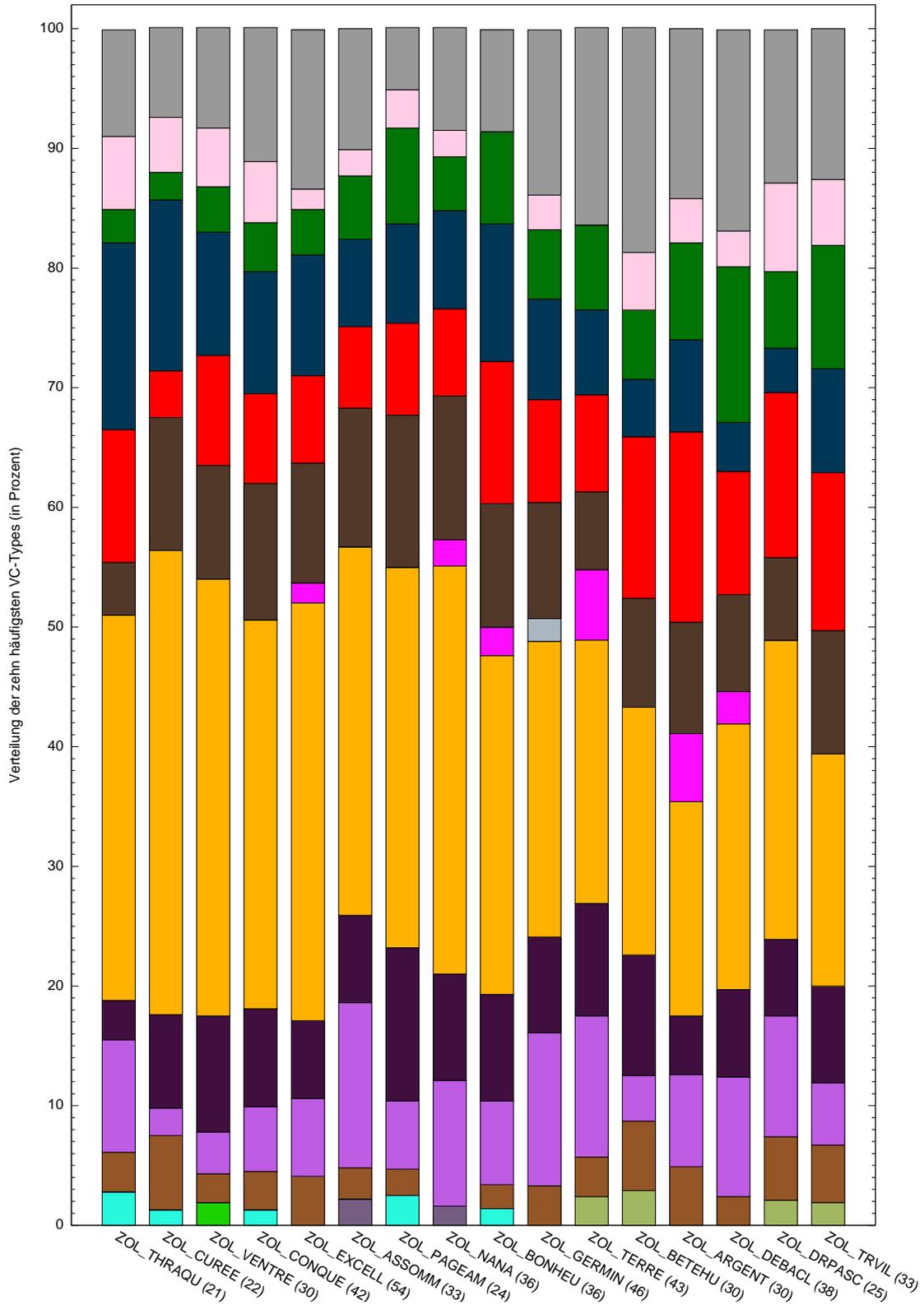


Abbildung 4.12: Die zehn häufigsten VC-Types innerhalb der Redebegleitung bei **Émile Zola** in prozentualer Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Types an. Die Romankürzel sind nach Abb. 3.1 aufzulösen.

(1580) – *Tu t'en fous, toi! tu manges!* hurle *Levaque*, en montrant le poing à *Rasseneur*.⁶⁹

Ein von anderen Autoren häufiger eingesetztes Kommunikationsverb, *ajouter*, gehört bei Zola nicht zu den bevorzugten VC-Types. Nur in einem einzigen Roman, *Le ventre de Paris* (1873), gehört es mit 1,9% zu den zehn häufigsten Verben. Gleiches gilt für *faire*. Es ist in keinem der 16 Romane in der Liste der zehn häufigsten VC-Types und ist mit Anteilen zwischen 0,1% (*Au bonheur des dames* (1883)) und 1,6% (*Le docteur Pascal* (1893)) nur mit wenigen Tokens vertreten. In insgesamt sechs Romanen wurde es überhaupt nicht ermittelt.

Émile Zola setzt wie die Autoren der zuvor analysierten Gruppen das Verbenpaar *s'écrier/crier* ein. War jedoch dort bislang immer *s'écrier* deutlich in der Überzahl, so hat sich dieses Verhältnis bei Zola ins Gegenteil verkehrt: *crier* erreicht in insgesamt sechs Romanen zweistellige Werte bis maximal 13,8% (*L'assommoir* (1877)) und wird damit zum standardisierten Ausdrucksmittel der in Zolas Romanen porträtierten einfachen Leute. In *Au bonheur des dames* (1883) und *La terre* (1887) ist *s'écrier* nicht einmal mehr unter den zehn häufigsten Kommunikationsverben anzutreffen und gehört mit Anteilen zwischen 1,1% (*Au bonheur des dames* (1883)) und 7,4% (*Le docteur Pascal* (1893)) zu den schwächer frequenten Kommunikationsverben.

Acht der 16 untersuchten Romane haben eine Gesamttokenzahl von 150 000 Tokens und mehr. Für alle diese Texte gilt jedoch, dass sie nur sehr wenig direkte Rede enthalten, d.h. der Anteil beträgt nur zwischen 11,9% (*Thérèse Raquin* (1867)) und 20,9% (*L'argent* 1891)). Somit hat der längste Roman sogar den geringsten Anteil von Figurenrede, und auch nur jedem dritten Turn (31%) folgt eine Redebegleitung. Angesichts dieser Werte ist die Zahl der VC-Types für die Redebegleitung mit 38 dennoch recht hoch. Vergleicht man jedoch diese Angaben mit den Werten in *Son Excellence Eugène Rougon* (1876), so fällt die Bilanz besser aus. Mit 136 916 Tokens ist der Text rund ein Drittel kürzer, enthält mehr als doppelt so viel direkte Rede (32,9%) und weist außerdem nicht nur einen höheren Anteil von redebegleitenden Sätzen (47%) auf, sondern erreicht mit 54 Types sogar das Maximum an verwendeten Verben. Die Anzahl der VC-Types innerhalb der Redebegleitung korreliert nicht zwangsläufig mit der Gesamtlänge des Textes oder dem Anteil der direkten Rede, sondern scheint in manchen Fällen vom Anteil des redebegleitenden *discours citant* abzuhängen. Zwei Beispiele bestätigen diese Annahme ebenfalls: Mit 123 600 Tokens gehört *La conquête de Plassans* (1874) zu den kürzeren Texten, hat aber von allen 16 analysierten Texten mit 41,1% den höchsten

⁶⁹ Émile Zola, *Germinal* (1885).

Anteil von direkter Rede, deren Turns zu 67% eine Redebegleitung enthalten. Gleichzeitig sind darin mit 42 VC-Types viele verschiedene Kommunikationsverben enthalten. *Germinal* (1885) wiederum ist mit 177 544 Tokens der drittlängste der 16 Romane, hat aber mit einem Anteil von 15,4% wenig direkte Rede, die aber zu 41% aus redebegleitenden Turns besteht. Die Zahl der VC-Types ist mit 46 ebenfalls sehr hoch. Gegen diese Annahme sprechen jedoch die Zahlen für *Une page d'amour* (1878). Mit einer Länge von 109 900 Tokens handelt es sich um einen der kürzeren Texte, der zu gut einem Fünftel (20,4%) aus direkter Rede besteht. Jeder zweite Turn (52%) steht in Verbindung mit einer Redebegleitung, und dennoch ist die Anzahl der VC-Types mit 24 Verben gering. Hier muss es noch einen anderen Faktor als die genannten geben, weshalb die Redebegleitung bei ähnlichen Voraussetzungen nicht so ausdifferenziert ist wie in anderen Romanen.⁷⁰

Auf die über die zehn häufigsten VC-Types hinaus vorhandenen Kommunikationsverben verteilen sich bei Émile Zola insgesamt zu wenige VC-Tokens, als dass eines davon durchschnittlich mehr als drei Tokens an sich binden könnte. Mit hochfrequenten Verben innerhalb der restlichen VC-Types ist daher auch bei Zola nicht zu rechnen.

Henri Alain-Fournier, Marcel Prévost, Marcel Proust und Colette

Bei diesen vier Autoren bewegt sich die Zahl der vorkommenden VC-Types zwischen 15 (Henri Alain-Fournier, *Le grand meaulnes* (1913)) und 67 (Colette, *Chéri* (1920)). Die Gesamtlänge der sieben betrachteten Texte variiert sehr stark. Während die Romane von Marcel Prévost, Henri Alain-Fournier und Colette deutlich unter 100 000 Tokens liegen – *Chéri* (1920) hat insgesamt sogar nur 41 235 Tokens –, sind die drei untersuchten Texte von Marcel Proust mit zwischen 181 851 (*Du côté de chez Swann* (1913)) und 247 669 Tokens (*Du côté de Guermantes* (1922)) sehr lang. Daraus sollte jedoch nicht geschlossen werden, dass aufgrund der Textlänge besonders viele VC-Types enthalten sind. Die meisten VC-Types, 67 an der Zahl, wurden wider Erwarten im kürzesten Text, *Chéri* (1920) von Colette, gezählt. Dies ist wohl vor allem der Tatsache geschuldet, dass *Chéri* zu 70,8% aus direkter Rede besteht, und auch wenn sich der Anteil der Turns mit redebegleitendem *discours citant* nur auf 26% beläuft, so werden für eine abwechslungsreiche Redebegleitung entsprechend viele Kommunikationsverben benötigt. Der Anteil der direkten Rede beträgt bei Marcel Proust nur zwischen 19,2% (*Du côté de chez Swann* (1913)) und 28,7% (*Du côté de Guermantes* (1922)), doch trotz höherer An-

⁷⁰ Ähnlich verhält es sich auch mit *Thérèse Raquin* (1867): 70 946 Tokens / 11,9% direkte Rede / 68% Redebegleitung / 21 VC-Types und *La curée* (1871): 113 110 Tokens / 12,1 % direkte Rede / 55% Redebegleitung / 22 VC-Types.

teile von redebegleitenden Sätzen (28-43%) sind die VC-Types mit zwischen 25 und 41 Verben weniger vielfältig. Dass *Le grand meaulnes* von Henri Alain-Fournier (1913) die wenigstens VC-Types aufweist, verwundert nicht weiter, da es sich um einen kurzen Text (68 607 Tokens) mit wenig direkter Rede (11,5%) handelt. Zwar ist zu 42% eine Redebegleitung vorhanden, doch dieser Umstand allein bewirkt offenbar keine automatische Zunahme der enthaltenen Kommunikationsverben.

Acht Kommunikationsverben kommen in allen untersuchten Romanen vor – *continuer, demander, dire, faire, interrompre, reprendre, répliquer* und *répondre* –, jedoch gehören diese nicht automatisch zu den häufigsten Verben eines jeden Textes. Weiterhin sind folgende Verben in fast allen Romanen unterschiedlich oft vorhanden: *ajouter, crier, déclarer, s'écrier, murmurer, penser, poursuivre* und *répéter*.

Abbildung 4.13 spiegelt sehr deutlich die unterschiedliche Gestaltung des *discours citant* durch die vier Autoren wider. Dies drückt sich unter anderem durch eine besonders große Vereinigungsmenge von 23 Verben aus. Trotz aller individuellen Unterschiede ist auch in dieser Gruppe *dire* das häufigste Kommunikationsverb. Seine Anteile liegen zwischen 19,7% (Marcel Prévost, *Les demi-vierges* (1894)) und 53,1% (Henri Alain-Fournier, *Le grand meaulnes* (1913)). Marcel Prévost setzt *dire* mit 19,7% (*Les demi-vierges* (1894)) und 25% (*L'automne d'une femme* (1893)) am sparsamsten ein.

Anders als die drei anderen Autoren dieser Gruppe scheint Marcel Prévost eine ausgeprägte Vorliebe für *faire* als Kommunikationsverb zu haben, da er dieses mit 19,4% (*Les demi-vierges* (1894)) und 20,8% (*L'automne d'une femme* (1893)) fast gleichberechtigt zu *dire* verwendet.⁷¹ Bis auf Anteile von 5,5% in *Le grand meaulnes* (1913) von Henri Alain-Fournier und 2,3% in *À l'ombre des jeunes filles en fleurs* (1918) von Marcel Proust sind die Vorkommen von *faire* in den übrigen Romanen quasi unbedeutend.⁷²

Neben dem hochfrequenten Verb *faire* gibt es bei Marcel Prévost weitere Auffälligkeiten. Verglichen mit den anderen fünf Romanen kommen in seinen beiden Texten die Verben *murmurer*⁷³ und *répliquer*⁷⁴ sehr häufig in der Redebegleitung vor. Das fast in allen Texten des Arbeitskorpus hochfrequente Verb *crier* wurde für Prévosts Romane überhaupt nicht ermittelt, während *s'écrier* nur in *Les demi-vierges* (1894) mit 2,3% zu den häufigsten VC-Types zählt und in *L'automne d'une femme* (1893) zwar vorkommt,

⁷¹ Ein höherer Anteil für *faire* innerhalb des gesamten Arbeitskorpus wurde nur noch in *Germinie Lacerteux* (1865) mit 23,8% ermittelt.

⁷² Die Werte erreichen für die verbleibenden Romane nicht einmal ein Prozent.

⁷³ *Les demi-vierges* (1894): 4,9%, *L'automne d'une femme* (1893): 6,7%. Die restlichen Texte haben für *murmurer* Anteile von lediglich 0,5% bis zu 1,3%.

⁷⁴ *Les demi-vierges* (1894): 13,5%, *L'automne d'une femme* (1893): 7%. Die Werte für die übrigen Texte bewegen sich für *répliquer* zwischen 0,3% und 0,8%.

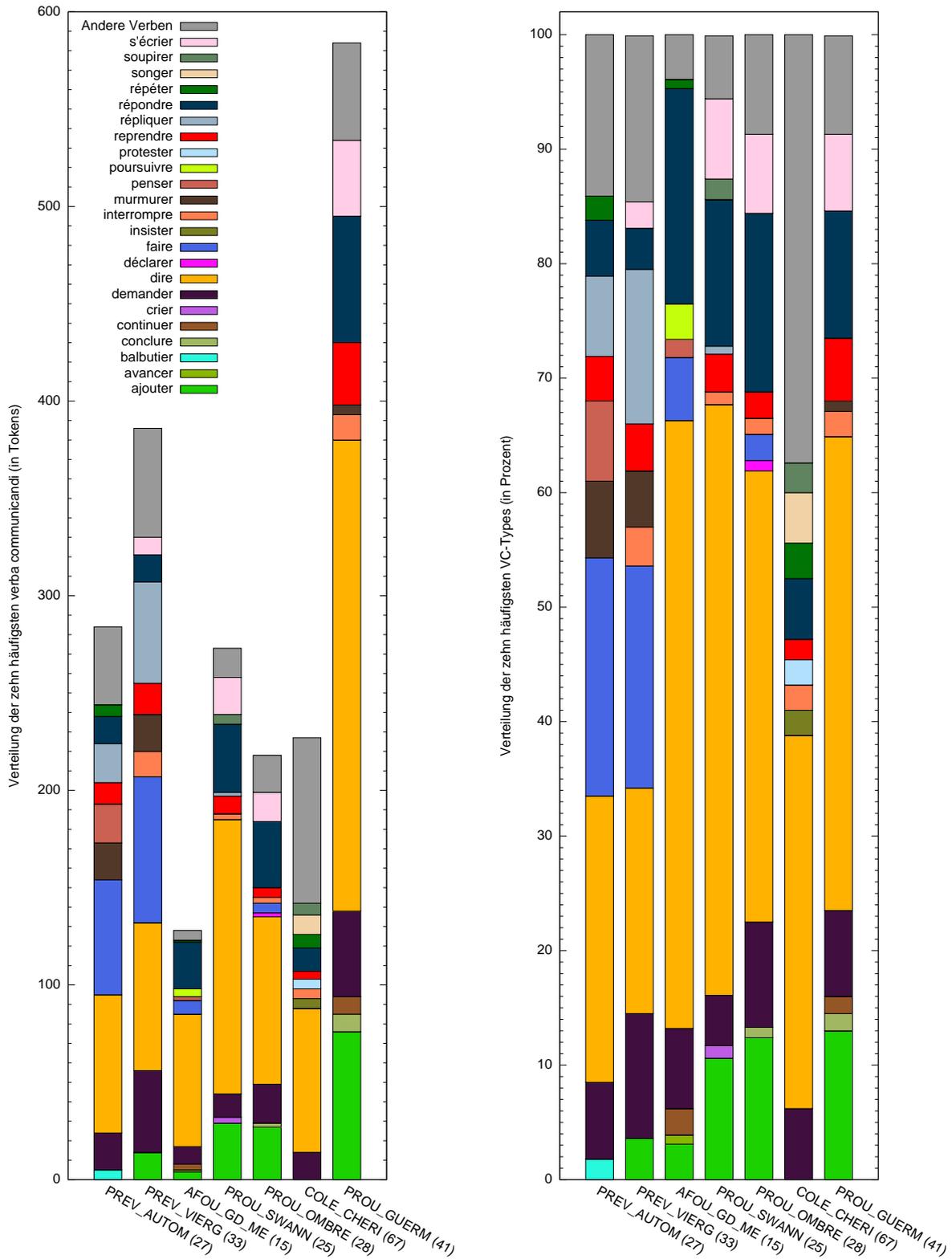


Abbildung 4.13: Die zehn häufigsten VC-Typen innerhalb der Redebegleitung bei **Marcel Prévost, Henri Alain-Fournier, Marcel Proust** und **Colette** in prozentualer Darstellung. Entlang der horizontalen Achse sind die einzelnen Romane chronologisch angeordnet. Die Zahl in Klammern hinter jedem Roman gibt die Gesamtzahl der VC-Typen an. Die Roman-kürzel sind nach Abb. 3.1 aufzulösen.

mit nur 1,8% der Gruppe der restlichen VC-Types zuzuordnen ist. In *L'automne d'une femme* werden die Gedanken der Figuren mit einem Anteil von 7% außerdem so oft wie in keinem anderen Text mittels direkter Rede ausgedrückt.⁷⁵ Nur bei Prévost kommt *balbutier* als Kommunikationsverb vor, das jedoch nur in *L'automne d'une femme* mit 1,8% Teil der zehn häufigsten VC-Types ist. In *Les demi-vierges* gehört auch *interrompre* mit 3,4% – dem Höchstwert innerhalb dieser Gruppe – zu den häufigsten VC-Types.

In *Le grand meaulnes* (1913) von Henri Alain-Fournier ist die Gestaltung der Redebegleitung aufgrund der wenigen VC-Types (15) und dem hohen Anteil von *dire* (53,1%) vergleichsweise unspektakulär, doch auch dieser Roman weicht hier und da von der bislang beobachteten Zusammenstellung der Kommunikationsverben ab. Neben *faire* (5,5%) wurden vor allem *ajouter* (3,1%), *poursuivre* (3,1%) und *continuer* (2,3%) häufiger ermittelt. Auch *penser* ist mit 1,8% noch Teil der häufigsten zehn VC-Types.

Für die Redebegleitung der drei untersuchten Romane von Marcel Proust ergibt sich ein vergleichsweise hoher Anteil von *dire*, der zwischen 39,4% (*À l'ombre des jeunes filles en fleurs* (1918)) und 51,6% (*Du côté de chez Swann* (1913)) liegt. Weitere höherfrequente Verben mit ungefähr gleichen Anteilen sind *ajouter* (10,7-13%), *répondre* (11,1-15,6%), *demander* (4,4-9,2%) und *s'écrier* (6,7-7%). In allen drei Texten nachgewiesen und zu den häufigsten zehn Verben gehörend, jedoch nicht in allzu großer Zahl vertreten sind außerdem *reprendre* und *interrompre*. Das vor allem bei Zola sehr beliebte Kommunikationsverb *crier* ist nur in *Du côté de chez Swann* Teil der häufigsten zehn VC-Types und mit 1,1% nur unbedeutend häufiger als in den beiden anderen Romanen. Die Verben *faire* und *déclarer* gehören nur in *À l'ombre des jeunes filles en fleurs* (1918)⁷⁶ gerade noch zu den häufigsten zehn VC-Types, während dies für *continuer* und *murmurer* nur in *Du côté de Guermantes* (1922)⁷⁷ der Fall ist.

Völlig anders gestaltet Colette ihre Redebegleitung in *Chéri* (1920). Ein gutes Drittel der Redebegleitung wird durch *dire* (32,6%) bestritten, gefolgt von drei weiteren, fast gleich häufig vorkommenden Verben: *demander* (6,2%), *répondre* (5,3%) und *songer* (4,4%). Letzteres ist als Kommunikationsverb in keinem anderen Text innerhalb des Arbeitskorpus häufiger vertreten als hier, so dass innerhalb der Redebegleitung zehn Turns ähnlich dem folgenden Beispiel ermittelt wurden:

(840) – *Mes dernières idylles, songeait Léa accoudée à la fenêtre.*

Die verbleibenden sechs der zehn häufigsten VC-Types kommen zwar deutlich seltener vor, dafür sind ihre Anteile fast gleich groß: *répéter* (3,1%), *soupirer* (2,6%), *insister*

⁷⁵ Für vier weitere Texte liegt der Anteil von *penser* zwischen 0,7% und 1,8%.

⁷⁶ Der Anteil von *faire* beläuft sich auf 2,3%, während der von *déclarer* nur 0,9% beträgt.

⁷⁷ Das Kommunikationsverb *continuer* hat hier einen Anteil von 1,5%, *murmurer* ist mit 0,9% vertreten.

(2,2%), *interrompre* (2,2%), *protester* (2,2%) und *reprendre* (1,8%). Augenfällig ist weiterhin, dass trotz der geringen Textlänge (41 235 Tokens) und einem Anteil von Redebegleitung von nur 26% weitere 57 VC-Types vorkommen, von denen der Großteil nur durch je ein VC-Token vertreten wird.

Für die restlichen VC-Types der übrigen Romane dieser Gruppe ergeben sich wie für den Großteil aller Texte keine Überraschungen, da zum einen viel weniger VC-Types über die zehn häufigsten hinaus eingesetzt werden und auch die VC-Tokens nicht so zahlreich sind, dass jeder VC-Type durch viel mehr als ein VC-Token repräsentiert wird.

4.2.2 Verteilung und Entwicklung der Konstruktionstypen

Die Gestaltungsmöglichkeiten für den redebegleitenden *discours citant* sind zwar vielfältig, aber aufgrund der Inversion von Subjekt und Verb auch nicht unendlich. Das hat zur Folge, dass bestimmte Konstruktionen immer wiederkehren, die je nach Autor bzw. Zeitpunkt der Publikation bezüglich ihrer Häufigkeit mehr oder weniger stark variieren. Bei den Konstruktionstypen, die zuvor in Kapitel 2.2.3 ausführlich vorgestellt und erläutert wurden, handelt es sich um die folgenden:

- **T1** / BASISTYP:
« ... », *dit-il.*
- **T2a** / KONSTRUKTION MIT ADJEKTIV ODER *participe passé*:
« ... », *répliqua-t-elle stupéfaite.*
- **T2b** / KONSTRUKTION MIT ADJEKTIV ODER *participe présent*:
« ... », *dit la jeune fille trouvant enfin la force de parler dans son émotion.*
- **T2c** / RELATIVSATZ-KONSTRUKTION:
« ... », *dit une voix qui ne lui était pas inconnue.*
- **T3a** / ADVERBIALKONSTRUKTION:
« ... », *s'écria vivement le jeune homme.*
- **T3b** / *gérondif*-KONSTRUKTION:
« ... », *remarqua-t-il en souriant.*
- **T3c** / PRÄPOSITIONALSYNTAGMA:
« ... », *répondit-elle avec quelque impatience.*
- **T3e** / KONSTRUKTION MIT DATIV-ERGÄNZUNG:
« ... », *murmura-t-elle à son mari.*
- **Toreille** / KONSTRUKTION MIT OREILLE:
« ... », *dit l'oncle à l'oreille du neveu.*
- **Tvoix** / KONSTRUKTION MIT VOIX/TON/AIR:
« ... »? *demanda-t-elle d' une voix altérée.*
- **TKon** / KONSTRUKTION MIT KONJUNKTION:
« ... », *dit César quand il la vit endormie.*

Die prozentualen Werte für jeden Roman und jeden Konstruktionstyp können in Anhang C in Tabelle C.1 eingesehen werden. Eine chronologische Übersichtsgaphik für das gesamte Arbeitskorpus bietet Abbildung 4.14. Die Verteilung der sechs häufigsten Konstruktionstypen wird in den Abbildungen 4.15 bis 4.20 noch einmal separat illustriert. Die Zahlen für die einzelnen Autoren und Romane fallen sehr unterschiedlich aus, und selbst in verschiedenen Werken ein und desselben Autors kommt es mitunter zu großen Schwankungen. Eine Aussage trifft jedoch auf alle Texte zu: Innerhalb des redebegleitenden *discours citant* dominiert stets die einfache Redebegleitung, d.h. der Basistyp T1 mit eingeschobenen Sätzen wie z.B. «dit-il». In mehr als der Hälfte aller Korpustexte – das sind 46 Romane – hat T1 einen Anteil von mehr als 50%. So beträgt der Anteil von T1 in 16 Romanen mehr als 50%, in 17 Romanen mehr als 60%, in acht Romanen mehr als 70% und in immerhin fünf Romanen sogar mehr als 80%.⁷⁸ Besteht die Redebegleitung zu mehr als 70% oder 80% aus dem Basistyp, was immerhin in acht bzw. fünf Romanen der Fall ist, so bleibt für die Verwendung weiterer Konstruktionen nur wenig Spielraum. Liegt der Anteil von T1 hingegen unter 50% – das trifft auf insgesamt 24 Texte des Arbeitskorpus zu –, darf man davon ausgehen, dass die Redebegleitung in diesen Texten deutlich elaborierter ist.⁷⁹

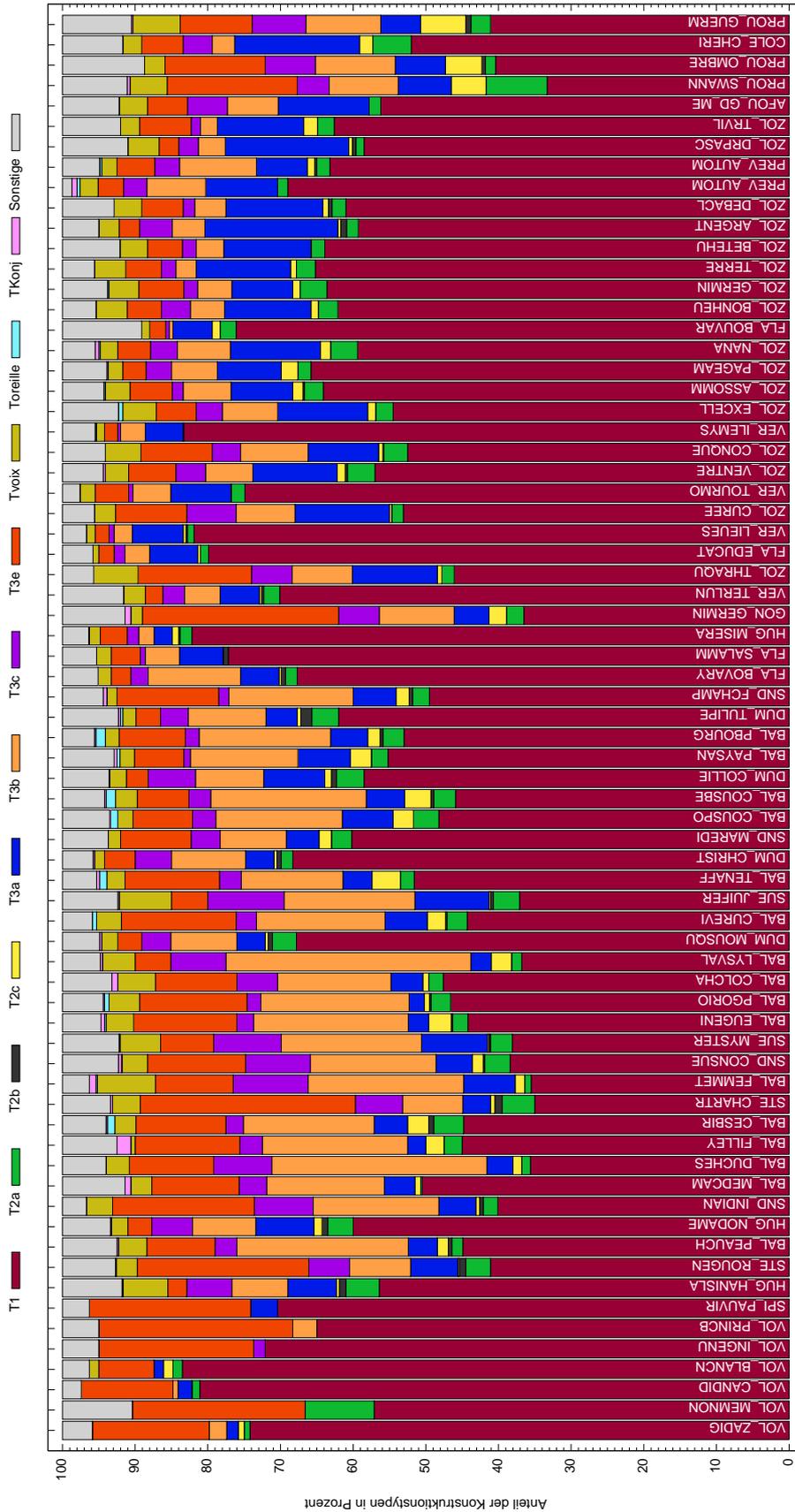
In einzelnen Fällen gibt es zwar Konstruktionstypen, die fast ebenso häufig zum Einsatz kommen wie T1, doch bilden diese die absolute Ausnahme.⁸⁰ Weitere prominente Konstruktionstypen, die in jedem Roman anzutreffen sind – wenn auch mit zuweilen stark voneinander abweichender Frequenz – sind die Adverbial- (T3a) und die *gérondif*-Konstruktion (T3b), die Konstruktion mit Präpositionalsyntax (T3c) sowie die Dativkonstruktion (T3e). Der Konstruktionstyp T2a (Konstruktion mit Adjektiv oder *participe passé*) gehört zu den selteneren, da seine Frequenz nur zwischen 0% und 9,4% liegt.⁸¹ Eine meist untergeordnete Rolle spielen die Konstruktionen mit Adjektiv oder *participe présent* (T2b), Konstruktionen mit *oreille* (Toreille) sowie die Konjunktionalkonstruktionen

⁷⁸ Die genauen Prozentwerte sind der Tabelle C.1 in Anhang C zu entnehmen.

⁷⁹ Von den 24 Romanen liegt der Anteil von T1 in 15 Fällen unter 50%, in neun unter 40%. Der geringste Anteil von T1 beträgt 33,3% und wurde in *Du côté de chez Swann* (1913) von Marcel Proust ermittelt.

⁸⁰ Dies trifft auf zwei Romane von Honoré de Balzac zu: *Le lys dans la vallée* (1844) und *La duchesse de Langeais* (1835). In ersterem wurden für T1 36,8% und für die *gérondif*-Konstruktion (T3b) 33,7% ermittelt. Für letzteren wurden für T1 35,6% sowie 29,6% für T3b gezählt. Ähnlich verhält es sich mit *La chartreuse de Parme* (1839) von Stendhal, wo T1 zu 35% und Dativkonstruktionen (T3e) zu 29,6% innerhalb der redebegleitenden Konstruktionen vorkommen. *Germinie Lacerteux* (1865) der Gebrüder de Goncourt enthält zu 36,5% den Basistyp und zu 27% Dativkonstruktionen.

⁸¹ Dieser Höchstwert wird in *Memnon* (1750) von Voltaire erreicht – dies ist um so erstaunlicher, als vor allem die frühen Texte kaum Variantenreichtum innerhalb der Redebegleitung aufweisen –, während die zweit- und dritthöchsten Werte von Texten am Ende des Beobachtungszeitraum erzielt werden, d.h. 8,4% für *Du côté de chez Swann* (1913) von Marcel Proust und 5,3% für *Chéri* (1920) von Colette. Für die übrigen Romane liegt der Anteil bei 4,6% oder darunter.



Romane in chronologischer Reihenfolge

Abbildung 4.14: Kompakte Übersicht über das gesamte Arbeitskorpus mit allen innerhalb der Redebegleitung vorkommenden Konstruktionstypen.

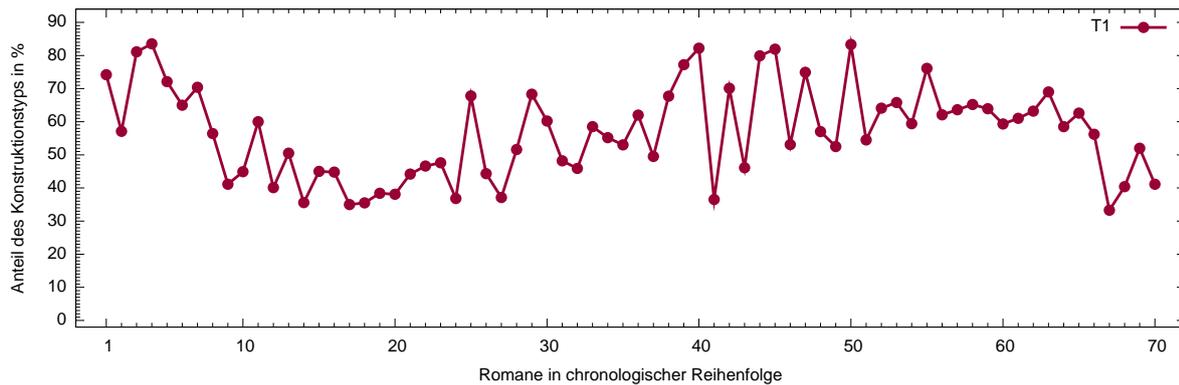


Abbildung 4.15: Konstruktionstyp T1 (« ... », *dit-il*.) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Obwohl er gewissen Schwankungen unterliegt, ist dies mit Abstand der häufigste Konstruktionstyp.

(TKonj), da sie allesamt nur selten die Ein-Prozent-Marke überschreiten.

Bei näherer Betrachtung der Abbildung 4.14, die die innerhalb der Redebegeleitung üblichen Konstruktionstypen illustriert, sind aufgrund der chronologischen Anordnung der Korpus-texte drei Phasen erkennbar, deren Merkmale im Folgenden erläutert werden.

Phase 1: 1747-1787 bzw. Texte Nr. 1-7⁸²

Die erste Phase umfasst den Zeitraum 1747 bis 1787, in den die Schaffenszeit der beiden Autoren Voltaire und Bernardin de Saint-Pierre fällt. Dass Voltaire oder Bernardin de Saint-Pierre hinsichtlich der Ausgestaltung der direkten Rede und des *discours citant* noch bei weitem nicht so kreativ waren wie später publizierende Autoren, wurde bereits durch die geringe Anzahl der verwendeten *verba communicandi* (s. Kapitel 4.2.1) deutlich. Dieser Eindruck bestätigt sich, wenn man die Verteilung der Konstruktionstypen innerhalb der redebegleitenden Sätze betrachtet: Der *discours citant* wird in dieser Gruppe zu mindestens 57% durch T1, «*dit-il*», bestritten, das Maximum beträgt 83,5% (*Le blanc et le noir* (1764)) bzw. 81,1% (*Candide* (1759)). Auch zu einem deutlich späteren Zeitpunkt liegt der Anteil in *Paul et Virginie* (1787) von Bernardin de Saint-Pierre noch bei 70,4%.

Neben T1 ist sowohl bei Voltaire als auch bei de Saint-Pierre die Dativkonstruktion (T3e) mit Werten zwischen 7,6% (*Le blanc et le noir* (1764)) und 26,7% (*La princesse de Babylone* (1774)) in größerer Zahl vertreten. Für T2a (attributive Ergänzung des Kommunikationsverbs) ist nur ein einziger Text, *Memnon* (1750), mit einem Anteil von 9,5% erwähnenswert, da in drei Romanen diese Konstruktion gar nicht enthalten ist

⁸² Die Aufschlüsselung der Romannummern erfolgt über Tabelle 3.1.

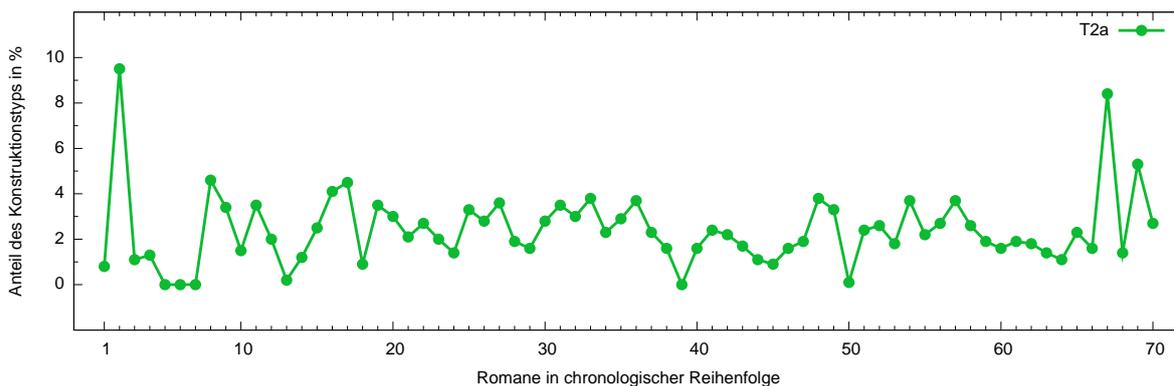


Abbildung 4.16: Konstruktionstyp T2a (« ... », *répliqua-t-elle stupéfaite.*) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Schwachfrequenter Konstruktionstyp, dessen Anteil mehrheitlich zwischen 2% und 4% liegt.

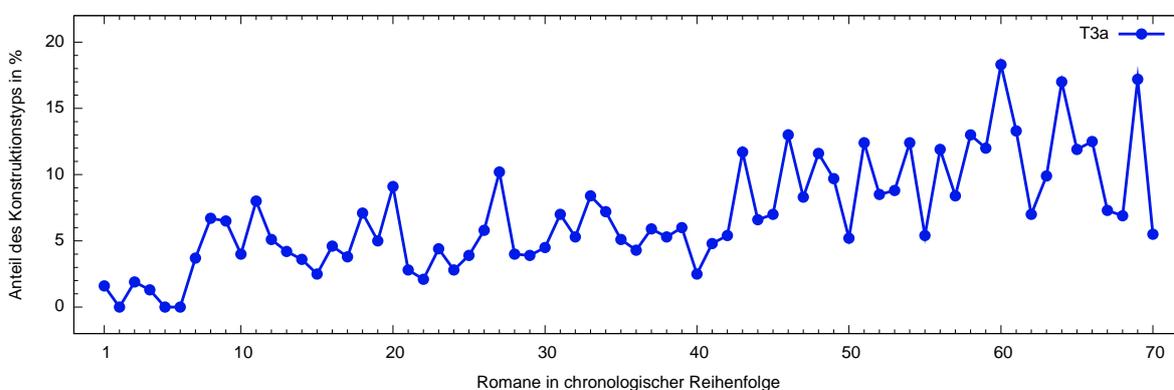


Abbildung 4.17: Konstruktionstyp T3a (« ... », *s'écria vivement le jeune homme.*) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Der Adverbialkonstruktion wird ab Mitte des Untersuchungszeitraums häufiger verwendet und erreicht Anteile von bis zu 18,3% (in *L'argent* (1891) von Émile Zola).

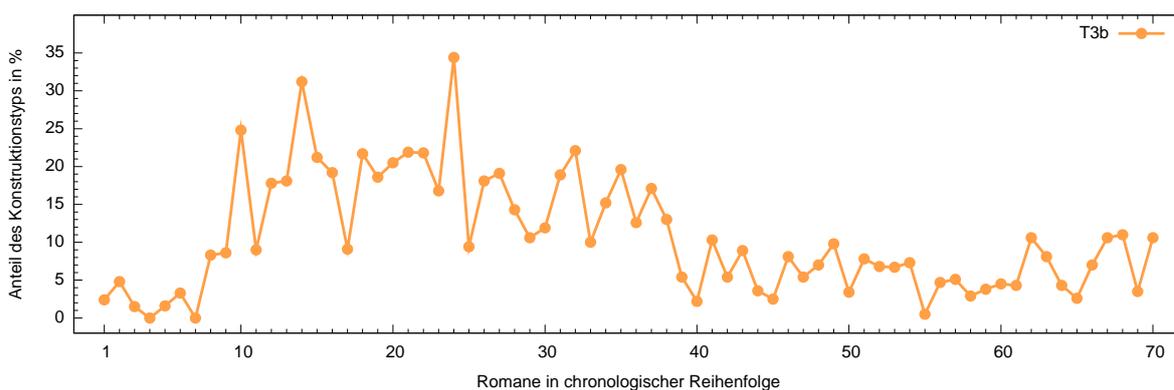


Abbildung 4.18: Konstruktionstyp T3b (« ... », *remarquait-il en souriant.*) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Die *gérondif*-Konstruktion erfreut sich vor allem in der ersten Korpuschälfte größerer Beliebtheit, die ab der zweiten Hälfte wieder deutlich nachlässt.

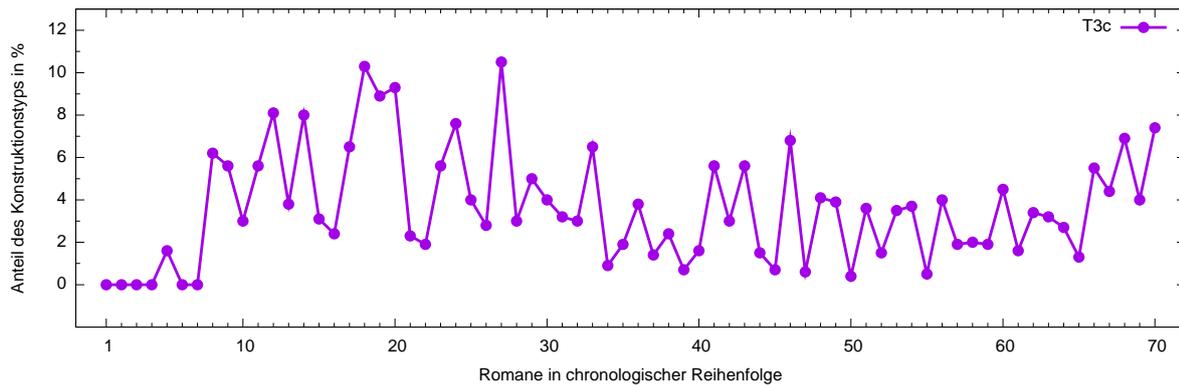


Abbildung 4.19: Konstruktionstyp **T3c** (« ... », *répondit-elle avec quelque impatience.*) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Die Präpositionalsyntagmen sind für die ersten Jahre des Untersuchungszeitraums gar nicht nachgewiesen und erreichen erst danach Anteile von bis zu 11%.

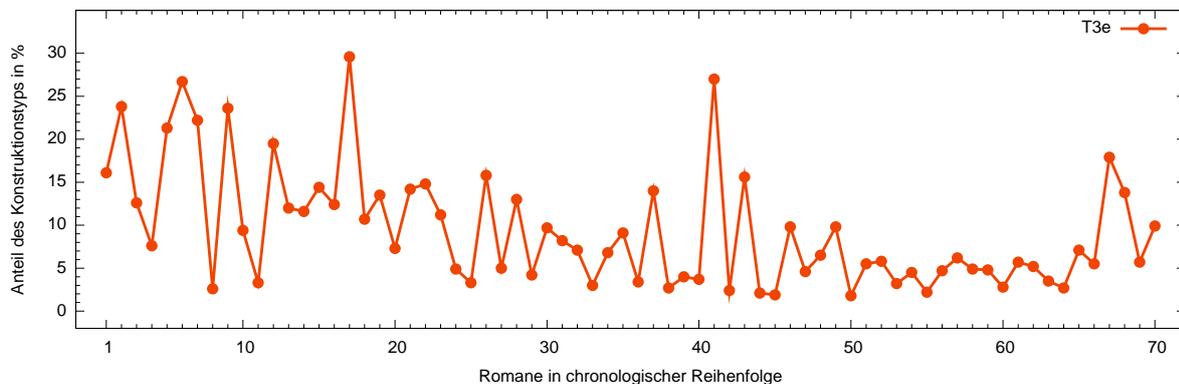


Abbildung 4.20: Konstruktionstyp **T3e** (« ... », *dit-elle à son mari.*) innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern (horizontale Achse) erfolgt über Tabelle 3.1. Die Dativkonstruktion ist vor allem in der ersten Korpushälfte häufig, bevor sie gegen Ende des Untersuchungszeitraums meist deutlich unter 10% sinkt.

und die drei restlichen Texte mit Werten zwischen 0,8% (*Zadig* (1747)) und 1,3% (*Le blanc et le noir* (1764)) diesbezüglich nicht ins Gewicht fallen. Alle weiteren untersuchten Konstruktionstypen kommen, sofern sie überhaupt nachgewiesen wurden, ebenfalls nur in geringer Zahl vor. Hervorzuheben ist dabei der Anteil von 3,7% für die Adverbialkonstruktion (T3a) in *Paul et Virginie* (1787) von Bernardin de Saint-Pierre sowie die Anteile für die *gérondif*-Konstruktion (T3b) in Voltaires *Zadig* (1747) und *La princesse de Babylone* (1768) mit 2,4% bzw 3,3%.

Die für Voltaire und Bernardin de Saint-Pierre stellenweise hohen Prozentwerte dürfen jedoch nicht darüber hinwegtäuschen, dass die Texte der beiden Autoren zum einen sehr viel kürzer sind als die übrigen im Korpus und nur zu maximal 33,9% (in *Candi-*

de (1759) von Voltaire) direkte Rede enthalten. Wenn daher in *Memnon* (1750) für T2a (Konstruktion mit Adjektiv oder *participe passé*) ein Anteil von 9,5% enthalten ist⁸³, so scheint dies auf den ersten Blick viel zu sein, tatsächlich wurde diese Konstruktion jedoch nur ganze zweimal ermittelt.

Phase 2: 1823-1850 bzw. Texte Nr. 8-37⁸⁴

Die zweite Phase erstreckt sich über den Zeitraum von 1823 bis 1850 und umfasst 30 Werke von Victor Hugo, Stendhal, George Sand, Honoré de Balzac und Eugène Sue.

Spätestens mit dem Erscheinen von *Han d'Islande* (1823) von Victor Hugo hat sich die Gestaltung des *discours citant* in der Romanliteratur grundlegend verändert. Zwar repräsentiert der Basistyp T1 («dit-il») mit zwischen 35% (Stendhal, *La chartreuse de Parme* (1839)) und 68,3% (Alexandre Dumas, *Le comte de Monte Christo* (1846)) immer noch die stärkste Konstruktion, aber die Redebegleitung ist insgesamt variantenreicher geworden. Neben der schon bei Voltaire üblichen Dativkonstruktion (T3e) kommen nun auch vermehrt Konstruktionen mit *gérondif* (T3b), mit Adjektiv oder *participe passé* (T2a) bzw. *participe présent* (T2b) sowie Adverbialkonstruktionen (T3a) und Präpositionalsyntagmen (T3c) vor. Besonders auffällig innerhalb dieser Phase ist die Vielzahl der Romane, deren redebegleitender *discours citant* zu weniger als 50%⁸⁵ bzw. 40%⁸⁶ aus dem Basistyp T1 besteht. Aus dem (vorübergehenden) Rückgang dieser Konstruktion resultiert eine Zunahme vor allem der *gérondif*-Konstruktion (T3b), der Adverbialkonstruktion (T3a), dem Präpositionalsyntagma (T3c) und der Relativsatzkonstruktion (T2c). Auch die zuvor praktisch nicht existente Konstruktion mit *oreille* (*Toreille*) ist in jedem Roman mit mehr oder weniger großen Anteilen vorhanden.

Auch wenn der Basistyp T1 bei **Honoré de Balzac** wie bei allen anderen Autoren

⁸³ Dies entspricht dem Maximalwert für T2a innerhalb des gesamten Korpus. Der zweithöchste Anteil wurde in *Du côté de chez Swann* (1913) ermittelt und beträgt 8,4%. Die Konstruktion ist dort trotz eines geringeren Anteils von direkter Rede immerhin 23 Mal enthalten.

⁸⁴ Die Aufschlüsselung der Romannummern erfolgt über Tabelle 3.1.

⁸⁵ In insgesamt zwölf Romanen liegt der Anteil von T1 zwischen 40% und 50%: Stendhal, *Le rouge et le noir* (1830): 41,1%; Honoré de Balzac, *La peau de chagrin* (1831): 44,9%; George Sand, *Indiana* (1832): 40,1%; Honoré de Balzac, *La fille aux yeux d'or* (1835): 45%; César Birotteau (1837): 44,8%; Eugénie Grandet (1843): 44,2%; Le père Goriot (1843): 46,6%; Le colonel Chabert (1844): 47,6%; Le curé de village (1845): 44,3%; Le cousin Pons (1847): 48,2%; La cousine Bette (1848): 45,9% und George Sand, *François le Champi* (1850): 49,5%.

⁸⁶ Weniger als 40% beträgt T1 in sieben Romanen auf: Honoré de Balzac, *La duchesse de Langeais* (1835): 35,6%; Stendhal, *La chartreuse de Parme* (1839): 35%; Honoré de Balzac, *La femme de trente ans* (1842): 35,5%; George Sand, *Consuelo* (1842): 38,4%; Eugène Sue, *Les mystères de Paris* (1842): 38,1%; Honoré de Balzac, *Le lys dans la vallée* (1844): 36,8% und Eugène Sue, *Le juif errant* (1845): 37,1%.

mit Anteilen zwischen 35,5%⁸⁷ und 55,2%⁸⁸ überwiegt, so ist dieser Romancier doch ein ausgesprochener Freund der *gérondif*-Konstruktion. In seinen Romanen liegen die Anteile für diesen Typ zwischen 14% (*Une ténébreuse affaire* (1846)) und 33,7% (*Le lys dans la vallée* (1844)) – so hoch wie bei kaum einem der anderen Autoren.⁸⁹ Damit liegt in letzterem Text der Anteil von T3b fast gleichauf mit dem Basistyp T1 (36,8%). Ähnlich nah beieinander liegen T1 und T3b außerdem in *La duchesse de Langeais* (1835) mit 35,6% bzw. 29,6%. Ab 1845 ist bei Balzac wieder eine leichte Abnahme der *gérondif*-Konstruktion feststellbar, welche mit der ebenfalls ab diesem Zeitpunkt einsetzenden Zunahme von T1 korreliert. Honoré de Balzac setzt neben T3b auch bevorzugt die Dativkonstruktion (T3e) ein, deren Werte zwischen 4,9%⁹⁰ und 15,8%⁹¹ und bis 1845 bis auf zwei Ausnahmen⁹² über 10% liegen. Nach 1845 ist ähnlich wie bei T3b ein Rückgang der Dativkonstruktionen zu beobachten – wohl auch hier wieder zugunsten des Basistyps –, bis 1850 der geringste Wert, 6,8%⁹³, erreicht ist. Konstruktionen wie T2b (Konstruktion mit Adjektiv oder *participe présent*), *Toreille* (Konstruktionen mit *oreille*) und *TKonj* (Konstruktionen mit angeschlossenem Konjunktionalsatz) wurden zwar nachgewiesen, fallen aber praktisch nicht ins Gewicht, da sie nur selten überhaupt einen Anteil von einem Prozent erreichen. Weniger häufig, dafür aber recht konstant vertreten, ist T3a (Adverbialkonstruktion) mit Anteilen zwischen 2,1%⁹⁴ und 7,2%⁹⁵. Während fast alle Typen außer des Basistyps ab 1845 eher seltener nachgewiesen sind, ist für T3a ab diesem Zeitpunkt sogar ein leichter Anstieg zu verzeichnen. Auch T3c (Präpositional-syntagma) ist mit Werten zwischen 0,9%⁹⁶ und 10,3%⁹⁷ für alle Romane nachgewiesen, allerdings nehmen die Werte dafür nach 1845 deutlich ab. Auch *Tvoix* ist bei Balzac mit Werten um 3% gut vertreten, und in *La femme de trente ans* (1842) wird mit 8% der Höchstwert für das gesamte Korpus erreicht.

Bei **Victor Hugo** besteht die Redebegleitung in allen drei untersuchten Romanen zu mehr als 50% aus T1. Für *Han d'Islande* (1823) und *Notre-Dame de Paris* (1831)

⁸⁷ Honoré de Balzac, *La femme de trente ans* (1842).

⁸⁸ Honoré de Balzac, *Les paysans* (1850).

⁸⁹ Innerhalb von Phase 2 liegt der Anteil von T3b nur in fünf Romanen zweier anderer Autoren über dem Minimum von 14% bei Balzac: in *Indiana* (1832): 17,3%, *Consuelo* (1842): 17,3% und *François le Champi* (1850): 17,1% von George Sand sowie in *Les mystères de Paris* (1842): 19,3% und *Le juif errant* (1845): 18% von Eugène Sue.

⁹⁰ Honoré de Balzac, *Le lys dans la vallée* (1844).

⁹¹ Honoré de Balzac, *Le curé de village* (1845).

⁹² Honoré de Balzac, *La peau de chagrin* (1831): 9,4% und *Le lys dans la vallée* (1844): 4,9%.

⁹³ Honoré de Balzac, *Les paysans* (1850).

⁹⁴ Honoré de Balzac, *Le père Goriot* (1843).

⁹⁵ Honoré de Balzac, *Les paysans* (1850).

⁹⁶ Honoré de Balzac, *Les paysans* (1850).

⁹⁷ Honoré de Balzac, *La femme de trente ans* (1842).

ist die Verteilung nahezu gleich, d.h. 56,4% bzw. 60% des redebegleitenden *discours citant* werden durch den Basistyp gebildet, während 7,7% bzw. 8,7% eine *gérondif*-Konstruktion (T3b) aufweisen, direkt gefolgt von der Adverbialkonstruktion (T3a), die mit 6,7% bzw. 8% ähnlich stark vertreten ist. Präpositionalsyntagmen (T3c) wurden in 6,2% bzw. 5,6% der Fälle ermittelt. Der Roman *Les misérables*, der 30 Jahre später erscheint, weist hingegen eine ganz andere Verteilung auf: Zu 82,2% handelt es sich innerhalb der Redebegleitung um T1, während der in seinen beiden anderen Romanen zweitstärkste Typ, die *gérondif*-Konstruktion (T3b), nur noch 2,1% erreicht. Auch die übrigen Konstruktionstypen sind nur mit maximal 3,7% (Dativergänzung, T3e) enthalten. Die in den frühen Romanen von Victor Hugo angetroffene Vielfalt innerhalb der Redebegleitung ist in *Les misérables* (1862) nicht mehr vorhanden.

Für die beiden untersuchten Romane von **Stendhal** gleicht die Verteilung der Konstruktionstypen in etwa der der beiden frühen Romane von Victor Hugo. Auf *Le rouge et le noir* (1830) und *La chartreuse de Parme* (1839) entfallen 41,1% bzw. 35% auf T1, gefolgt von 23,6% bzw. 29,6% für die Dativkonstruktion (T3e). Der Rest verteilt sich im Wesentlichen auf T3b (*gérondif*-Konstruktion) mit 8,4% bzw. 8,3%, T3c (Präpositionalsyntagma) mit 5,6% bzw. 6,5%, T3a (Adverbialkonstruktion) mit 6,5% bzw. 3,8%. Zu den seltener vertretenen Konstruktionen gehören schließlich T2a (Konstruktion mit Adjektiv oder *participe passé*) mit 3,4% bzw. 4,5% sowie Tvoix mit 2,9% bzw. 3,8%.

Die Werke von **Eugène Sue** und **George Sand** fallen vor allem durch einen vergleichsweise hohen Anteil von *gérondif*-Konstruktionen (T3b) auf, der sich in allen Romanen bis auf *La mare au diable* (1846)⁹⁸ von George Sand zwischen 17,3%⁹⁹ und 19,3%¹⁰⁰ bewegt. Trotzdem dominiert T1 mit Werten zwischen 37,1%¹⁰¹ und 60,2%¹⁰². Recht häufig ist auch T3c (Präpositionalsyntagma) mit Werten zwischen 8,1%¹⁰³ und 10,5%¹⁰⁴, allerdings trifft dies nicht auf die beiden späteren Romane *La mare au diable* (1846) und *François le champi* (1850) von George Sand zu, für die jeweils nur 4% bzw. 1,4% erreicht werden. Mit 9,1% und 10,2% sind in *Les mystères de Paris* (1842) und *Le juif errant* (1845) rund doppelt so viele Adverbialkonstruktionen (T3a) enthalten wie in den Romanen von George Sand. In diesen ist wiederum die Dativkonstruktionen (T3e)

⁹⁸ Die *gérondif*-Konstruktion ist hier nur mit 9,1% nachgewiesen.

⁹⁹ George Sand, *Indiana* (1832).

¹⁰⁰ Eugène Sue, *Les mystères de Paris* (1842).

¹⁰¹ Eugène Sue, *Le juif errant* (1845).

¹⁰² George Sand, *La mare au diable* (1846). Dieser hohe Wert für T1 («dit-il») geht einher mit einem geringen Anteil der *gérondif*-Konstruktion (T3b: 9,1%), die in ihren anderen Werken viel häufiger nachgewiesen wurde.

¹⁰³ George Sand, *Indiana* (1832).

¹⁰⁴ Eugène Sue, *Le juif errant* (1845).

mit Anteilen zwischen 9,7%¹⁰⁵ und 19,5%¹⁰⁶ sehr viel stärker vertreten als bei Eugène Sue. Während Eugène Sue den Konstruktionstyp T_{voix} immerhin noch mit 5,6%¹⁰⁷ und 7,2%¹⁰⁸ einsetzt, liegt dieser bei George Sand nur zwischen 1,4%¹⁰⁹ und 3,6%¹¹⁰. Kaum zum Einsatz kommt bei beiden Autoren Typ T_{2a}, d.h. die Konstruktion mit Adjektiv oder *participle passé*.

Phase 3: 1857-1913 bzw. Texte 38-70¹¹¹

Die dritte Phase umfasst den Zeitraum 1857-1913, innerhalb dessen 33 Romane der Autoren Gustave Flaubert, Victor Hugo, Jules und Édmond de Goncourt, Jules Verne, Émile Zola, Marcel Prévost, Henri Alain-Fournier, Colette und Marcel Proust entstanden sind.

Charakteristisch für diese Phase ist der wieder deutlich angestiegene Anteil von T₁, der in 27 von 29 Romanen¹¹² 50% und mehr beträgt. In acht Romanen¹¹³ beläuft sich T₁ auf mehr als 50%, in elf¹¹⁴ auf mehr als 60%, in fünf¹¹⁵ auf mehr als 70%, und in drei Texten¹¹⁶ besteht die Redebegleitung sogar zu mehr als 80% aus diesem Basistyp. Das hat zur Folge, dass die übrigen Konstruktionstypen meist nur noch eine untergeordnete Rolle spielen. Im Unterschied zu Phase 2, innerhalb derer die *gérondif*-Konstruktion (T_{3b}) nicht nur bei Honoré de Balzac sehr populär ist, ist T_{3b} zwar auch in dieser dritten

¹⁰⁵ George Sand, *La mare au diable* (1846).

¹⁰⁶ George Sand, *Indiana* (1832).

¹⁰⁷ Eugène Sue, *Les mystères de Paris* (1842).

¹⁰⁸ Eugène Sue, *Le juif errant* (1845).

¹⁰⁹ George Sand, *François le Champi* (1850).

¹¹⁰ George Sand, *Indiana* (1832).

¹¹¹ Die Aufschlüsselung der Romannummern erfolgt über Tabelle 3.1.

¹¹² Bei den beiden Ausnahmen handelt es sich um den Roman *Germinie Lacerteux* (1865) der Gebrüder de Goncourt, in dem T₁ nur zu 36,5% vertreten ist sowie um *Thérèse Raquin* (1867) von Émile Zola, wo T₁ in 46,1% des redebegleitenden *discours citant* anzutreffen ist.

¹¹³ Dies trifft auf sieben Texte von Émile Zola zu: *La curée* (1871): 53,1%, *Le ventre de Paris* (1873): 57%, *La conquête de Plassans* (1874): 52,5%, *Son Excellence Eugène Rougon* (1876): 54,5%, *Nana* (1879): 59,4%, *L'argent* (1891): 59,3% und *Le docteur Pascal* (1893): 58,5%. 56,2% wurden außerdem in *Le grand meaulnes* (1913) von Henri Alain-Fournier erreicht.

¹¹⁴ Der Konstruktionstyp T₁ erreicht mehr als 60% in *Madame Bovary* (1857) von Gustave Flaubert: 67,7%, in acht Romanen von Émile Zola: *L'assommoir* (1877): 64,1%, *Une page d'amour* (1878): 65,8%, *Au bonheur des dames* (1883): 62,1%, *Germinal* (1888): 63,3%, *La terre* (1887): 65,2%, *La bête humaine* (1890): 63,9%, *La débâcle* (1892): 61%, *Les trois villes* (1894): 62,6% sowie in *L'automne d'une femme* (1893): 62,6% und *Les demi-vierges* (1894): 62,6% von Marcel Prévost.

¹¹⁵ Zu diesen gehören Gustave Flauberts *Salammbô* (1862): 77,2%, *L'éducation sentimentale* (1869): 79,9%, *Bouvard et Pécuchet* (1881): 76,1% und *De la terre à la lune* (1865): 70,1%, *Le tour du monde en quatre-vingt jours* (1873): 74,9%, beide von Jules Verne.

¹¹⁶ Dies ist der Fall in *Les misérables* (1862) von Victor Hugo (82,2%) und in den beiden Romanen von Jules Verne, *Ving-mille lieues sous les mers* (1869): 81,9% und *L'île mystérieuse* (1875): 83,3%.

Phase vertreten, aber bei weitem nicht mehr so häufig, denn nur in fünf Fällen erreicht dieser Typ überhaupt zweistellige Anteilwerte¹¹⁷. Auch der Anteil der Dativkonstruktion (T3e) hat abgenommen. Obwohl *Germinie Lacerteux* (1865) von Jules und Édmond de Goncourt mit dem Maximum von 27% herausragt¹¹⁸, beträgt der Anteil nur in vier Texten¹¹⁹ 10% oder mehr (in Phase 2 sind es noch 15 Texte). Ebenfalls seltener als bereits in Phase 1 ist die Relativsatzkonstruktion (T2c), da sie für die Mehrzahl der Texte maximal ein Prozent erreicht. Nur Marcel Proust verwendet die Konstruktion häufiger, weshalb die Werte zwischen 4,8%¹²⁰ und 6,2%¹²¹ liegen. Deutlich zugenommen hat hingegen der Anteil der Adverbialkonstruktion (T3a), die in immerhin 14 von 33 Romanen zweistellige Werte mit maximal 18,3%¹²² erreicht. In 17 weiteren Texten liegt der Anteil zwischen 5% und 10%, während er nur in zwei Werken¹²³ weniger als 5% beträgt.

In den vier Romanen von **Alexandre Dumas** ist T1 mit Werten zwischen 58,5%¹²⁴ und 68,3%¹²⁵ klar dominant. Der zweithäufigste Typ ist die *gérondif*-Konstruktion, deren Anteil jeweils zwischen 9,1%¹²⁶ und 10,7%¹²⁷ beträgt. Während die Adverbialkonstruktion (T3a) in *Le collier de la reine* (1849) mit 8,4% fast gleichaufliegt mit der *gérondif*-Konstruktion, ist ihr Anteil in den übrigen drei Romanen nur noch halb so groß. Der Anteil von Präpositionalsyntagmen (T3c) bewegt sich zwischen 3,8%¹²⁸ und 6,5%¹²⁹ und liegt damit knapp über den Werten für Dativkonstruktionen (T3e) und Konstruktionen mit Adjektiv oder *participe passé* (T2a). Während Tvoix noch Anteile zwischen 1,4%¹³⁰ und 2,3%¹³¹ erreicht, überschreiten die Werte für alle restlichen Konstruktionen nur in einem

¹¹⁷ Gustave Flaubert, *Madame Bovary* (1857): 12,7%, Jules und Édmond de Goncourt, *Germinie Lacerteux* (1865): 10,3%, Marcel Prévost, *L'automne d'une femme* (1893): 10,6%, Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918): 11% und *Le côté de Guermantes* (1922): 10,3%.

¹¹⁸ Innerhalb des Korpus wird dieser Anteil nur noch von *La chartreuse de Parme* (1839) von Stendhal mit 29,6% übertroffen

¹¹⁹ Jules und Édmond de Goncourt, *Germinie Lacerteux* (1865): 27%, Émile Zola, *Thérèse Raquin* (1867): 15,6%, Marcel Proust, *Du côté de chez Swann* (1913): 17,9% und *À l'ombre des jeunes filles en fleurs* (1918): 13,8%.

¹²⁰ Marcel Proust, *Du côté de chez Swann* (1913).

¹²¹ Marcel Proust, *Le côté de Guermantes* (1922).

¹²² Émile Zola, *L'argent* (1891).

¹²³ Victor Hugo, *Les misérables* (1862): 2,5% und Jules und Édmond de Goncourt, *Germinie Lacerteux* (1865): 4,8%.

¹²⁴ Alexandre Dumas, *Le collier de la reine* (1849).

¹²⁵ Alexandre Dumas, *Le comte de Monte Christo* (1846).

¹²⁶ Alexandre Dumas, *Les trois mousquetaires* (1844).

¹²⁷ Alexandre Dumas, *La tulipe noire* (1850).

¹²⁸ Alexandre Dumas, *La tulipe noire* (1850).

¹²⁹ Alexandre Dumas, *Le collier de la reine* (1849).

¹³⁰ Alexandre Dumas, *Le comte de Monte Christo* (1846).

¹³¹ Alexandre Dumas, *Le collier de la reine* (1849).

einzigsten Fall die Ein-Prozent-Marke¹³².

Vor allem die Werte für *Germinie Lacerteux* (1865) der **Gebrüder de Goncourt** fallen im Vergleich zu den anderen Romanen in dieser Gruppe auf, da der Anteil von T1 mit 36,5% so gering ist wie in keinem anderen Roman innerhalb des Korpus. Die Dativ-Konstruktion ist mit 27% fast ebenso häufig vertreten und wird innerhalb des Korpus nur noch von *La chartreuse de Parme* (1839) von Stendhal übertroffen. Die *gérondif*-Konstruktion ist die einzige weitere Gruppe, die mit 10,3% einen zweistelligen Wert aufweist. Auch Präpositionalsyntagmen (T3c) sind mit 5,6% recht häufig vertreten, ebenso die Adverbialkonstruktion (T3a) mit 4,8%, während (T2a) und Relativsatzkonstruktionen (T2c) jeweils nur auf 2,4% kommen. Das Schlusslicht bilden die Gruppen *Tvoix* und *TKonj* mit 1,6% bzw. 0,8%.

In den vier betrachteten Romanen von **Gustave Flaubert** macht T1 mit Werten zwischen 67,7%¹³³ und 79,9%¹³⁴ den Hauptanteil aus. Dieser Maximalwert wird nur noch in wenigen anderen Werken übertroffen¹³⁵. Andere Konstruktionstypen sind somit nur schwach ausgeprägt und erreichen nur in einem einzigen Fall überhaupt einen zweistelligen Wert¹³⁶. Die nach T1 stärksten Typen sind die Adverbialkonstruktion (T3a) und die *gérondif*-Konstruktion (T3b). Während T3a mit Werten zwischen 5,3%¹³⁷ und 6,6% fast konstant bleibt, ist für T3b innerhalb von 24 Jahren eine kontinuierliche Abnahme von anfänglich 12,7% auf gerade noch 0,5%¹³⁸ zu verzeichnen. Schwach vertreten sind Relativsatzkonstruktionen (T3c) oder Konstruktionen mit *voix/ton/oreille* (*Tvoix*). Die restlichen Typen, z.B. T2a, T2b, T2c, fallen mit maximal 2,2%¹³⁹ kaum ins Gewicht oder wurden, wie z.B. für *Toreille* und *TKonj*, überhaupt nicht nachgewiesen.

Die vier untersuchten Romane von **Jules Verne** fallen durch eine überdurchschnittlich häufige Verwendung von T1, dessen Werte sich zwischen 70,1%¹⁴⁰ und 83,3%¹⁴¹ bewegen, auf. Entsprechend niedrig, d.h. mit Werten im nur einstelligen Bereich, sind die Anteile anderer Konstruktionstypen. So erreicht der zweitstärkste Typ, die Adverbi-

¹³² Alexandre Dumas, *La tulipe noire* (1850).

¹³³ Gustave Flaubert, *Madame Bovary* (1857).

¹³⁴ Gustave Flaubert, *L'éducation sentimentale* (1869).

¹³⁵ Voltaire, *Candide* (1759): 81,1%, *Le blanc et le noir* (1764): 83,5%, Victor Hugo, *Les misérables* (1862): 82,2%, Jules Verne, *Ving-mille lieues sous les mers* (1869): 81,9% und *L'île mystérieuse* (1875): 83,3%.

¹³⁶ In *Madame Bovary* (1857) von Gustave Flaubert erreicht die *gérondif*-Konstruktion (T3b) einen Anteil von 12,7%.

¹³⁷ Gustave Flaubert, *Madame Bovary* (1857).

¹³⁸ Gustave Flaubert, *Bouvard et Pécuchet* (1881).

¹³⁹ Ermittelt für Typ 2a in *Bouvard et Pécuchet* (1881) von Gustave Flaubert.

¹⁴⁰ Jules Verne, *De la terre à la lune* (1865).

¹⁴¹ Jules Verne, *L'île mystérieuse* (1875).

alkonstruktion (T3a), Werte zwischen 5,2%¹⁴² und 8,3%¹⁴³. Am dritthäufigsten ist die *gérondif*-Konstruktion, deren Anteil zwischen 2,5%¹⁴⁴ und 5,2%¹⁴⁵ liegt. Der Höchstwert der Dativkonstruktion (T3e) fällt im Vergleich zu anderen Romanen innerhalb dieser Phase gering aus und liegt bei nur 4,6%¹⁴⁶. In den restlichen drei Texten wird die Redebegleitung nur zu einem sehr geringen Teil mittels Dativkonstruktion gebildet, d.h. die Anteilwerte liegen zwischen 2,4%¹⁴⁷ und 1,8%¹⁴⁸. Bei letzterem Wert handelt es sich um das korpusweite Minimum, d.h. in keinem anderen im Korpus enthaltenen Roman sind weniger Dativkonstruktionen enthalten. Geringfügig aus dem Rahmen fällt das Präpositionalsyntax (T3c) in *De la terre à la lune* (1865) mit 3%, während es in den übrigen drei Werken nicht einmal ein Prozent erreicht. Ähnlich verhält es sich auch mit dem Konstruktionstyp *Tvoix*, der hier ebenfalls mit 3% nachgewiesen wurde, während er in den übrigen Texten deutlich darunter liegt. Dies erklärt sich sicherlich dadurch, dass der früheste der vier untersuchten Verne-Romane mit 70,1% auch den geringsten Anteil von T1 aufweist und somit andere Konstruktionstypen etwas häufiger vorkommen können. Gar nicht oder nur bis zu maximal 0,4% kommen bei Jules Verne Konstruktionen mit Adjektiv oder *participe présent* (T2b), Relativsatzkonstruktionen (T2c), Konstruktionen mit *voix/ton/oreille* (*Toreille*) sowie Konstruktionen mit Konjunktion (*TKonj*) vor.

Bei **Émile Zola** wird der *discours citant* innerhalb der Redebegleitung überwiegend mithilfe von T1 und T3a (Adverbialkonstruktionen) bestritten. Bis auf *Thérèse Raquin* (1867), wo T1 nur zu 46,1% vorkommt, liegt der Anteil von T1 in den 15 anderen Texten zwischen mindestens 52,5%¹⁴⁹ und höchstens 65,8%¹⁵⁰. Nach 1877 beträgt der Anteil von T1 für elf Romane bis auf wenige Ausnahmen 60% oder mehr. Die Adverbialkonstruktion (T3a) ist die zweithäufigste. Ihre Werte erreichen zwischen 8,4%¹⁵¹ und 18,3%¹⁵² und liegen für 12 von 16 Texten im zweistelligen Bereich. Aufgrund des stetigen Auf und Ab ist jedoch keine besondere Entwicklung über die Zeit hinweg erkennbar. Ein wenig anders verhält es sich mit der Dativkonstruktion (T3e). Im ältesten der 16 untersuchten Romane, *Thérèse Raquin* (1867), ist der Anteil mit 15,6% noch zweistellig, während er danach auf unter 10% fällt und mit gewissen Schwankungen nach und

¹⁴² Jules Verne, *L'île mystérieuse* (1875).

¹⁴³ Jules Verne, *Le tour du monde en quatre-vingt jours* (1873).

¹⁴⁴ Jules Verne, *Vingt-mille lieues sous les mers* (1869).

¹⁴⁵ Jules Verne, *Le tour du monde en quatre-vingt jours* (1873).

¹⁴⁶ Jules Verne, *Le tour du monde en quatre-vingt jours* (1873).

¹⁴⁷ Jules Verne, *De la terre à la lune* (1865).

¹⁴⁸ Jules Verne, *L'île mystérieuse* (1875).

¹⁴⁹ Émile Zola, *La conquête de Plassans* (1874).

¹⁵⁰ Émile Zola, *Une page d'amour* (1878).

¹⁵¹ Émile Zola, *Germinal* (1885).

¹⁵² Émile Zola, *L'argent* (1891).

nach bis auf 2,7%¹⁵³ sinkt. Wie bei Jules Verne spielt auch bei Émile Zola die *gérondif*-Konstruktion (T3b) eine untergeordnete Rolle. Dabei gilt: Je stärker T1 vertreten ist, desto geringer fällt der Anteil von T3b aus. Bis zum Erscheinen von *Nana* im Jahre 1877 liegt T3b ohne große Abweichungen bei zwischen 6,3%¹⁵⁴ und 9,3%¹⁵⁵. Für die elf danach publizierten Romane ist eine stetige Abnahme zu beobachten, die schließlich mit 2,3% in *Les trois villes* (1894) ihren Tiefpunkt erreicht. Eine ähnliche Entwicklung ist für T3c (Präpositionalsyntaxma) festzustellen. Während die redebegleitenden Sätze in *Thérèse Raquin* (1867) und *La curée* (1871) noch einen Anteil von 5,6% bzw. 6,8% an Präpositionalsyntaxmen aufweisen, liegt dieser für die restlichen 14 Romane bei 4,5%¹⁵⁶ oder deutlich darunter. Recht häufig sind außerdem Konstruktionen mit *voix/ton/air* (Tvoix) enthalten, die Anteile zwischen 2%¹⁵⁷ und 6,1%¹⁵⁸ erreichen. Seltener kommen dagegen Konstruktionen mit Adjektiv oder *participe passé* (T2a)¹⁵⁹ und Relativsatzkonstruktionen¹⁶⁰ vor. Mit einem Anteil von weniger als einem Prozent oder gar nicht nachgewiesen sind jeweils Konstruktionen mit Adjektiv oder *participe présent* (T2b) sowie Konstruktionen mit *oreille* (Toreille) oder Konjunktionen (TKonj).

Die Werte für die beiden untersuchten Romane von **Marcel Prévost**, *L'automne d'une femme* (1893) und *Les demi-vierges* (1894), weichen nur minimal voneinander ab. Zum einen ist der Anteil von T1 mit 69%¹⁶¹ bzw. 63,2%¹⁶² wie in 14 anderen Romanen sehr hoch. Aufgrund der Dominanz von T1 bleibt wenig Spielraum für weitere Konstruktionstypen, so dass der restliche *discours citant* im Wesentlichen durch Adverbial- (9,9% bzw. 7%) und *gérondif*-Konstruktionen (8,1% bzw. 10,6%) gebildet wird. Auf die Dativkonstruktionen (T3e), Präpositionalsyntaxmen (T3c) und Konstruktionen mit *voix/ton/air* (Tvoix) entfällt mit Werten zwischen 2,1% und 5,2% noch ein geringer Teil, während andere Konstruktionstypen entweder noch seltener sind oder gar nicht vorkommen.

In *Le grand Meaulnes* (1913) von **Henri Alain-Fournier** wird mindestens jeder zweite redebegleitende *discours citant* durch Typ1, *dit-il*, realisiert. Die Adverbialkonstruktion (T3a) ist die einzige weitere Konstruktion mit einem Wert im zweistelligen Bereich, d.h. 12,5%. Die übrigen redebegleitenden Sätze setzen sich im Wesentlichen aus *gérondif*-

¹⁵³ Émile Zola, *Le docteur Pascal* (1893).

¹⁵⁴ Émile Zola, *Une page d'amour* (1878).

¹⁵⁵ Émile Zola, *La conquête de Plassans* (1874).

¹⁵⁶ Émile Zola, *L'argent* (1891).

¹⁵⁷ Émile Zola, *Une page d'amour* (1878).

¹⁵⁸ Émile Zola, *Thérèse Raquin* (1867).

¹⁵⁹ Die Anteile liegen zwischen 1,1% (*Le docteur Pascal* (1893)) und 3,8% (*Le ventre de Paris* (1873)).

¹⁶⁰ Diese sind mit Anteilen zwischen 0,3% (*La curée* (1871)) und 2,3% (*Une page d'amour* (1878)) vertreten.

¹⁶¹ Marcel Prévost, *L'automne d'une femme* (1893).

¹⁶² Marcel Prévost, *Les demi-vierges* (1894).

dif-Konstruktionen (T3b mit 7%), Präpositionalsyntagmen und Dativkonstruktionen (T3c bzw. T3e mit je 5,5%) zusammen.

Marcel Prousts Romane sind gekennzeichnet durch einen im Vergleich zu anderen Werken innerhalb dieser Phase geringen Anteil von T1. *Du côté de chez Swann* (1913) weist diesbezüglich mit 33,3% den geringsten Wert innerhalb des gesamten Korpus auf. Für *À l'ombre des jeunes filles en fleurs* (1918) und *Le côté de Guermantes* (1922) liegt der Anteil von T1 mit 40,4% bzw. 41,4% ein wenig höher. Eine weitere Vorliebe Prousts ist die häufiger als bei anderen Autoren anzutreffende Relativsatzkonstruktion (T2c) innerhalb des Redeeinleitung. Während diese allenfalls noch bei Honoré de Balzac häufiger zum Einsatz kommt und in keinem anderen Werk auch nur annähernd die Fünf-Prozent-Marke erreicht, liegt sie in *Du côté de chez Swann* (1913) immerhin bei 4,8%, in *À l'ombre des jeunes filles en fleurs* (1918) bei 5% und in *Le côté de Guermantes* (1922) sogar bei 6,2%. Für *Du côté de chez Swann* (1913) ist außerdem der Wert für die attributive Konstruktion (T2a) hervorzuheben, der sich mit 8,4% zwar unterhalb des Spitzenwertes von 9,5% für *Memnon* (1750)¹⁶³ bewegt, aber doch deutlich über dem in Colettes *Chérie* (1920) erreichten Anteil von 5,3% liegt. Während T2b und *TKonj* mit Werten bis zu 0,7% nur mit sehr wenigen Beispielen vorhanden sind, wurde *Toreille* gar nicht nachgewiesen.

In **Colettes** Roman *Chéri* (1920) ist mit 52% jeder zweite redebegleitende *discours citant* dem Typ T1 zuzuordnen. Während sich die Verteilung der Konstruktionstypen in diesem Text nicht nennenswert von denen der anderen Romane dieser Gruppe unterscheidet, fallen doch vor allem drei Werte auf: Zum einen kommen adverbiale Konstruktionen (T3a) überdurchschnittlich häufig vor, d.h. mit einem Anteil von 17,2% liegt *Chéri* auf das gesamte Korpus bezogen an zweiter Stelle. Nur *L'argent* (1891) von Émile Zola liegt noch 1,1% Prozentpunkte darüber. Dem gegenüber steht mit 3,1% ein wie auch bei Jules Verne und Émile Zola sehr geringer Anteil von *gérondif*-Konstruktionen (T3b). Colette setzt wie Marcel Proust die attributive Konstruktion (T2a) deutlich häufiger als andere Schriftsteller ein, weshalb T2a mit 5,3% korpusweit an zweiter Stelle nach Prousts *Du côté de chez Swann* (1913) liegt. Vergleichbar häufig ist auch die Dativkonstruktion (T3e), die einen Anteil von 5,7% erreicht. Weniger frequent sind hingegen das Präpositionalsyntagma T3c mit 4%, *Tvoix* mit 2,6% und die Relativsatzkonstruktion (T2c) mit 1,8%, während T2b, *Toreille* und *TKonj* überhaupt nicht nachgewiesen wurden.

¹⁶³ Man beachte, dass sich aufgrund der unterschiedliche Länge der beiden Texte bei fast gleicher prozentualer Häufigkeit völlig unterschiedliche absolute Zahlen ergeben: Der Anteil von 9,5% in Voltaires *Memnon* (1750) für T2a (Konstruktion mit Adjektiv oder *participe passé*) entspricht zwei Vorkommen, während T2a für *Du côté de chez Swann* (1913) von Marcel Proust mit 8,4% bzw. 23 Vorkommen nachgewiesen wurde.

4.2.3 Verteilung der VC-Types innerhalb der hochfrequenten Konstruktionstypen

Nachdem in Kapitel 4.2.2 erörtert wurde, welche Konstruktionstypen innerhalb der Redebegeleitung eingesetzt werden, lohnt sich zusätzlich ein Blick auf die Verteilung der VC-Types innerhalb der Konstruktionstypen, um festzustellen, welche *verba communicandi* wo besonders häufig vorkommen.

Die Abbildungen 4.21 bis 4.27 zeigen die Verteilung der häufigsten VC-Types innerhalb der hochfrequenten Konstruktionstypen. Die Darstellung der Werte erfolgt jeweils doppelt: Die linke Graphik bildet die ermittelten VC-Tokens in absoluten Zahlen ab, während die rechte zwecks besserer Vergleichbarkeit zwischen den Konstruktionstypen die gemittelten Werte in Prozent anzeigt.

Abbildung 4.21 liefert eine Gesamtdarstellung für alle im Korpus enthaltenen Texte. Daran schließen sich weitere Schaubilder für die zu Beginn von Kapitel 4 definierten Autorengruppen an, und zwar für Voltaire und Bernardin de Saint-Pierre (Abb. 4.22), für Victor Hugo, Stendhal, George Sand und Eugène Sue (Abb. 4.23), für Honoré de Balzac (Abb. 4.24), für Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert (Abb. 4.25), für Émile Zola (Abb. 4.26) und für Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette (Abb. 4.27). Für die häufigsten Konstruktionstypen T1, T2a, T3a, T3b, T3c und T3e¹⁶⁴ wurden jeweils die zehn häufigsten VC-Types ermittelt. Die Verben, die in der Legende abgebildet sind, stellen die Vereinigungsmenge dar. Die Größe der Vereinigungsmenge bewegt sich zwischen 11¹⁶⁵ und 22¹⁶⁶ Redeverben, was für insgesamt sieben betrachtete Autorengruppen einen Schnitt von knapp 14 Verben ergibt. In vier von sieben Schaubildern¹⁶⁷ umfasst die Vereinigungsmenge exakt 13 Verben. Ein *verbum communicandi*, das zwar für einen, nicht aber für einen anderen Konstruktionstyp abgebildet ist, kann dennoch in diesem vorkommen – es gehört dann nur nicht zu den zehn frequentesten Verben. Je größer die Vereinigungsmenge, desto stärker variieren die Verben von Konstruktionstyp zu Konstruktionstyp. Das gilt in besonderem Maße für die Gruppe Prévost/Alain-Fournier/Proust/Colette (Abb. 4.27). Da der Anteil des Basistyps T1 (*dit-il*) in 46 Texten 50% oder mehr beträgt und die Summe der darin enthaltenen VC-Tokens diejenige in anderen Konstruktionstypen um ein Vielfaches übertrifft, wurde die Anzahl der VC-Tokens für T1 zugunsten einer besseren Darstellung

¹⁶⁴ Auflösung der Konstruktionstypenkürzel mittels Übersicht auf S. 113.

¹⁶⁵ Dies gilt für die Texte von Voltaire und Bernardin de Saint-Pierre (siehe Abb. 4.22).

¹⁶⁶ Eine Vereinigungsmenge von 22 Redeverben wird in der Gruppe Prévost/Alain-Fournier/Proust/Colette erreicht (siehe Abb. 4.27).

¹⁶⁷ Siehe Abb. 4.21, Abb. 4.23, Abb. 4.25 und 4.26.

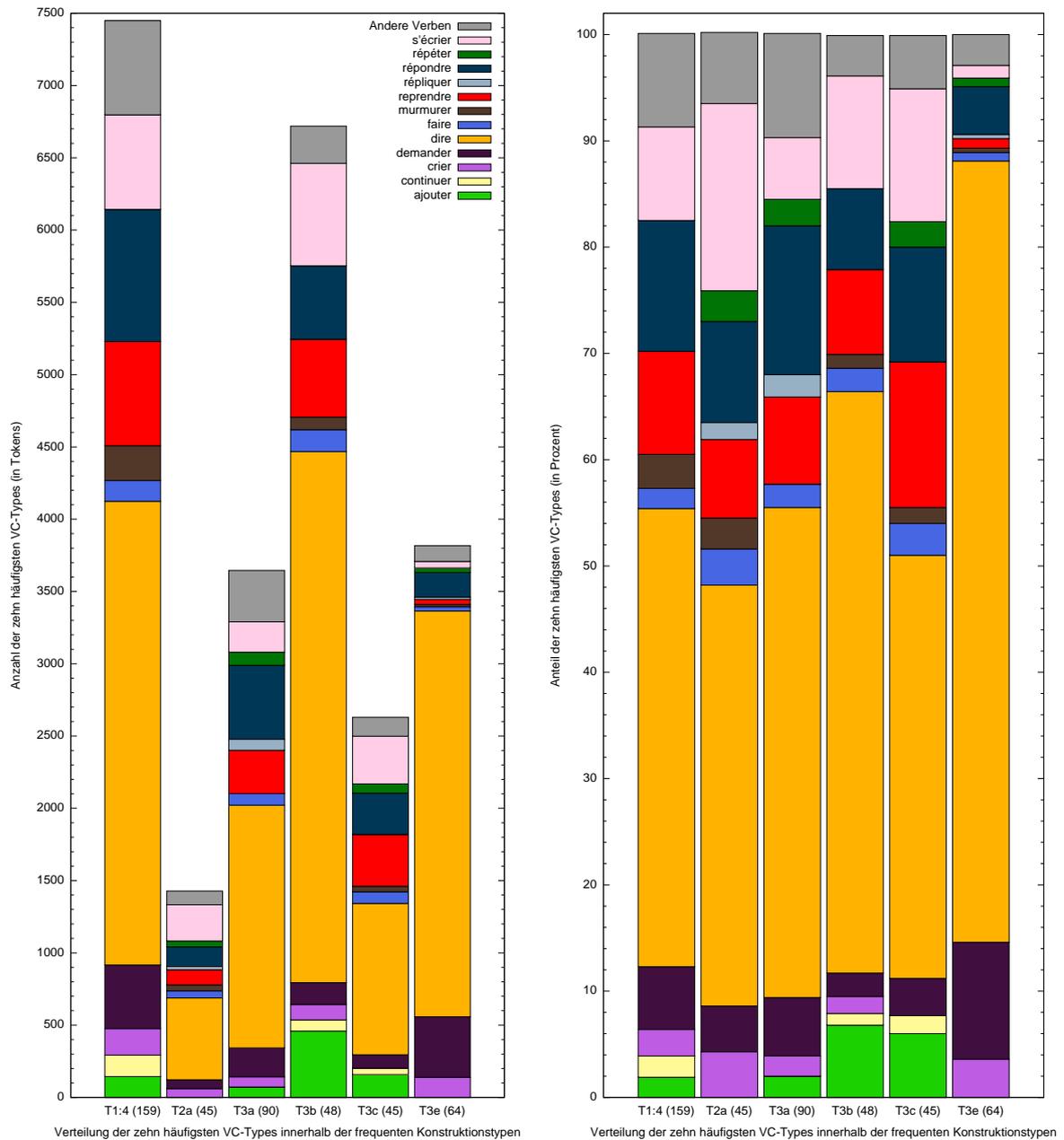


Abbildung 4.21: Verteilung der zehn häufigsten VC-Types pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) innerhalb des **gesamten Korpus**. Die senkrechte Achse gibt die Anzahl der VC-Tokens (linkes Bild) bzw. den Anteil in Prozent (rechtes Bild) wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Types an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

in allen Schaubildern geviertelt. Auf der horizontalen Achse gibt der Wert in Klammern hinter der Typenbezeichnung die Gesamtzahl der VC-Types an.

Die Auswertung der Schaubilder erfolgt nach zwei Gesichtspunkten: Erstens wird betrachtet, wie die Verteilung der *verba communicandi* für das gesamte Korpus aussieht. Einige Verben gehören in allen Konstruktionstypen zu den zehn häufigsten, andere hingegen sind nur in einzelnen Konstruktionstypen unter den höherfrequenten Verben zu finden. Zweitens ist zu überprüfen, in welchem Maße die auf das gesamte Korpus geltenden Aussagen auf die einzelnen Autorengruppen zutreffen.

Allen Autorengruppen gemein ist die Dominanz des Basistyps T1 (*dit-il*), dessen Anteil in mehr als der Hälfte der Romane, d.h. in 46 von 70 analysierten Texten, 50% oder mehr beträgt (siehe auch Abb. 4.15). Der Konstruktionstyp T1 weist sowohl für das gesamte Korpus als auch für jede der Autorengruppen die größte Vielfalt bezüglich der VC-Types auf.¹⁶⁸ Auch die Adverbial- (T3a) und die Dativkonstruktion (T3e) enthalten in allen Autorengruppen viele verschiedene Redeverben.¹⁶⁹ Dennoch lässt sich anhand der absoluten Anzahl der VC-Tokens nicht auf die Anzahl der VC-Types schließen, d.h. viele VC-Tokens innerhalb eines Konstruktionstyps sind kein Garant für eine entsprechende Vielzahl von VC-Types. Dies illustriert Abbildung 4.24 zu Balzac, in dessen Texten die *gérondif*-Konstruktion (T3b) nach T1 am stärksten vertreten ist, in der aber insgesamt nur 15 VC-Types vorkommen. Die Dativkonstruktion (T3e), die nur etwa halb so viele VC-Tokens aufweist, enthält hingegen 23 VC-Types, und auch die Adverbialkonstruktion (T3a), hat trotz einer noch niedrigeren Anzahl von VC-Tokens immerhin 21 VC-Types.¹⁷⁰

Das Universal-Redeverb *dire* (gelber Stapel) ist in allen besprochenen Konstruktionstypen das mit Abstand häufigste. Betrachtet man die Verteilung der *verba communicandi* für das gesamte Korpus in Abbildung 4.21, so gehören zu den zehn häufigsten Redeverben sechs Verben, die für alle Konstruktionstypen nachgewiesen wurden, und zwar *dire*, *demander*, *s'écrier*, *faire*, *répondre* und *reprendre*. Bestimmte *verba communicandi* treten wiederum nur in bestimmten Konstruktionstypen gehäuft auf. So ist beispielsweise *ajouter* in allen Konstruktionstypen bis auf T2a (Konstruktion mit Adjektiv oder *participe passé*) und T3e (Dativkonstruktion)¹⁷¹ frequent. Ähnlich verhält es sich für die

¹⁶⁸ In den Abbildungen gibt entlang der horizontalen Achse jeweils die Zahl in Klammern hinter der den Typenbezeichnung die Zahl der unterschiedlichen VC-Types an.

¹⁶⁹ In der Gruppe Prévost/Alain-Fournier/de Goncourt/Colette ändert sich diese Verteilung geringfügig, indem innerhalb der *gérondif*-Konstruktion (T3b) mehr Verben als innerhalb der Dativkonstruktion (T3e) enthalten sind.

¹⁷⁰ Deutliche Unterschiede gibt es auch in der Gruppe Hugo/Stendhal/Sand/Sue (4.23): Obwohl auch dort die *gérondif*-Konstruktion (T3b) fast doppelt so viele VC-Tokens wie die Adverbialkonstruktion (T3a) enthält, wurden nur halb so viele VC-Tokens (23) gezählt.

¹⁷¹ Für eine Dativkonstruktion müsste das *verbum communicandi* um ein indirektes Objekt erweitert wer-

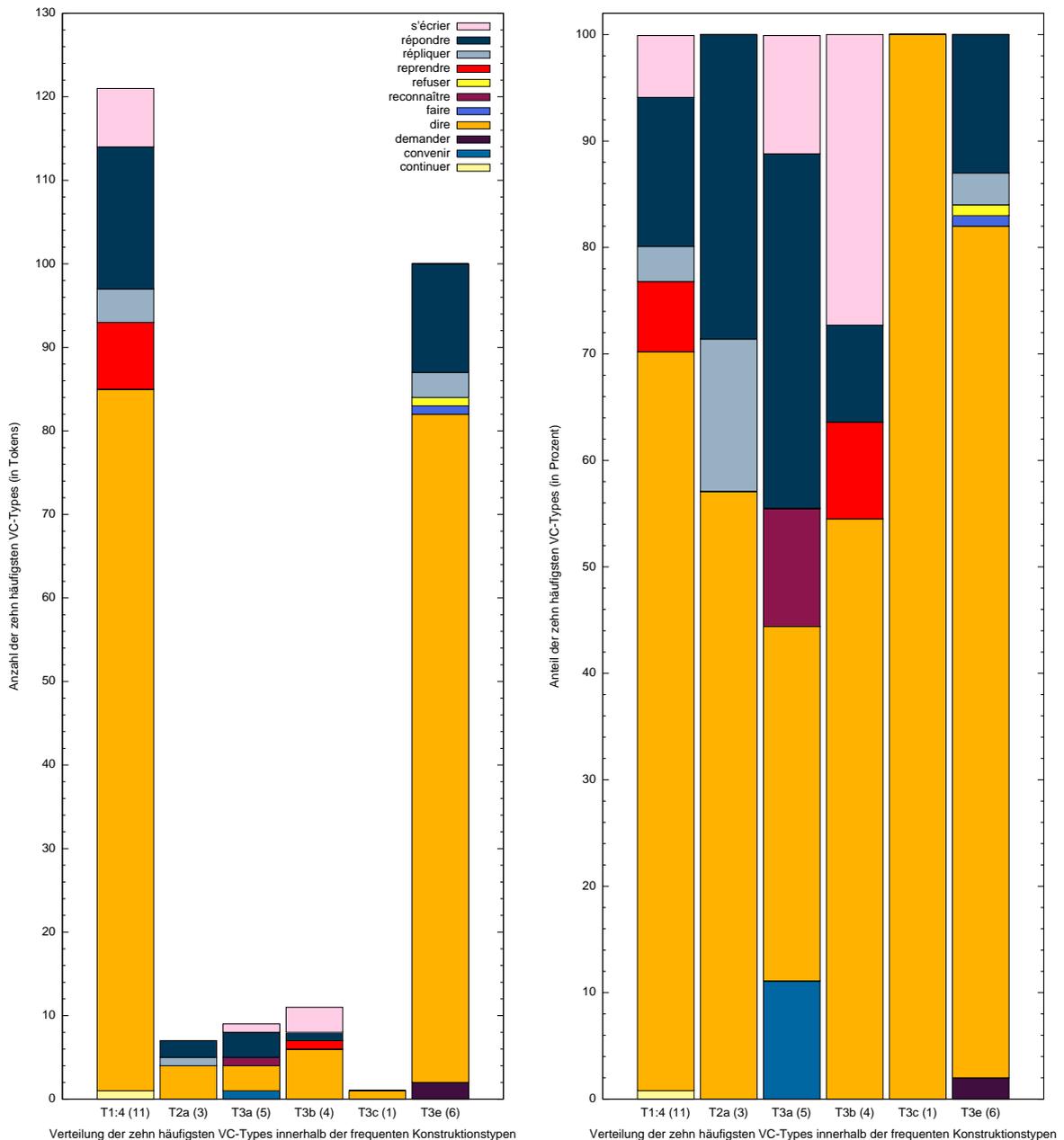


Abbildung 4.22: Verteilung der zehn häufigsten VC-Typen pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Voltaire** und **Bernardin de Saint-Pierre**. Die senkrechte Achse gibt die Anzahl der VC-Tokens (linkes Bild) bzw. den Anteil in Prozent (rechtes Bild) wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Typen an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

Verben *continuer*¹⁷², *crier*¹⁷³, *répéter*¹⁷⁴ und *répliquer*¹⁷⁵. Auf das gesamte Korpus bezogen befinden sich unter den zehn häufigsten Redeverben nur drei, die die Stimme der sprechenden Figur modifizieren: *crier*, *s'écrier* und *murmurer*. Zwar sind gerade in jüngeren Romanen viele weitere expressive Redeverben enthalten, jedoch sind in keiner der Autorengruppen mehr als vier¹⁷⁶ unter den häufigsten zehn.

Ein Verb, das zwar innerhalb des gesamten Korpus keine bedeutende Rolle spielt, weil es nur Anteile von maximal 2,7% erreicht, und zudem in drei Autorengruppen, d.h. Hugo/Stendhal/Sand/Sue sowie Balzac und Prévost/Alain-Fournier/Proust/Colette immer nur in T1 zu den häufigsten zehn Redeverben gehört, ist *penser*. Auch wenn sich die denkende Figur nicht aktiv in ein Gespräch einbringt, so erfährt der Leser dennoch mehr über sie durch Offenlegen ihres Innenleben.

Interessant ist die Verteilung der beiden Redeverben *s'écrier* und *crier*. Für das gesamte Korpus gilt, dass sich zwar beide Verben mit Ausnahme von T3c unter den zehn häufigsten *verba communicandi* befinden, *s'écrier* aber deutlich häufiger ist als *crier*. Dass *crier* bei Voltaire und Bernardin de Saint-Pierre noch gar nicht vorkommen, verwundert nicht weiter, denn das bewusste Einsetzen von Lautstärke¹⁷⁷ ziemt sich zu dieser Zeit weder für Adlige noch Gelehrte.

Auffällig ist außerdem das Ungleichgewicht in der Verwendung der beiden Verben *demander* und *répondre*. Innerhalb der Konstruktionstypen T1, T2a, T3a, T3b und T3c wurde *répondre* sehr viel häufiger ermittelt als *demander*. Die Fragehandlung kommt also in vielen Fällen anders als durch die Verwendung von *demander* zustande. So gehen Redebegleitungen mit *répondre* in vielen Fällen unbegleitete Turns voraus, in denen die Frage nur durch die Interpunktion ausgedrückt wird, wie z.B.

(247) – *Je vous approuve, dit la prisonnière d'une voix tremblante de douleur et de colère ; mais, ajouta-t-elle d'un ton fier, j'espère, seigneur Ordener, que vous ne m'avez pas entendue prier ?*

(248) – *Comtesse, répondit enfin le jeune homme, je vous ai entendue.*¹⁷⁸

Aber auch eine Kombination aus Interpunktion und Redeverb *dire* ist durchaus üblich:

den, was im Falle von *ajouter* jedoch nicht möglich ist.

¹⁷² *Continuer* befindet sich innerhalb von T1, T3b und T3c unter den zehn häufigsten Verben.

¹⁷³ *Crier* gehört in allen Konstruktionstypen bis auf T3c zu den häufigsten Redeverben.

¹⁷⁴ *Répéter* tritt in T2a, T3a, T3c und T3e gehäuft auf.

¹⁷⁵ Hochfrequent ist *répliquer* in T2a, T3a und T3e.

¹⁷⁶ Siehe Abbildung 4.26 zu Zola. In der Legende finden sich als expressive Verben *crier*, *s'écrier*, *murmurer* und *souffler*.

¹⁷⁷ Definitionen von *crier*: «Parler fort de manière à être entendu nettement et de loin, quand on exprime explicitement une chose d'importance.» (Quelle: TLFi, <http://atilf.atilf.fr/>, Eintrag unter B.2) bzw. «Dire d'une voix forte.» (Quelle: Rey-Debove & Rey (1993), S. 510).

¹⁷⁸ Victor Hugo, *Han d'Islande* (1823).

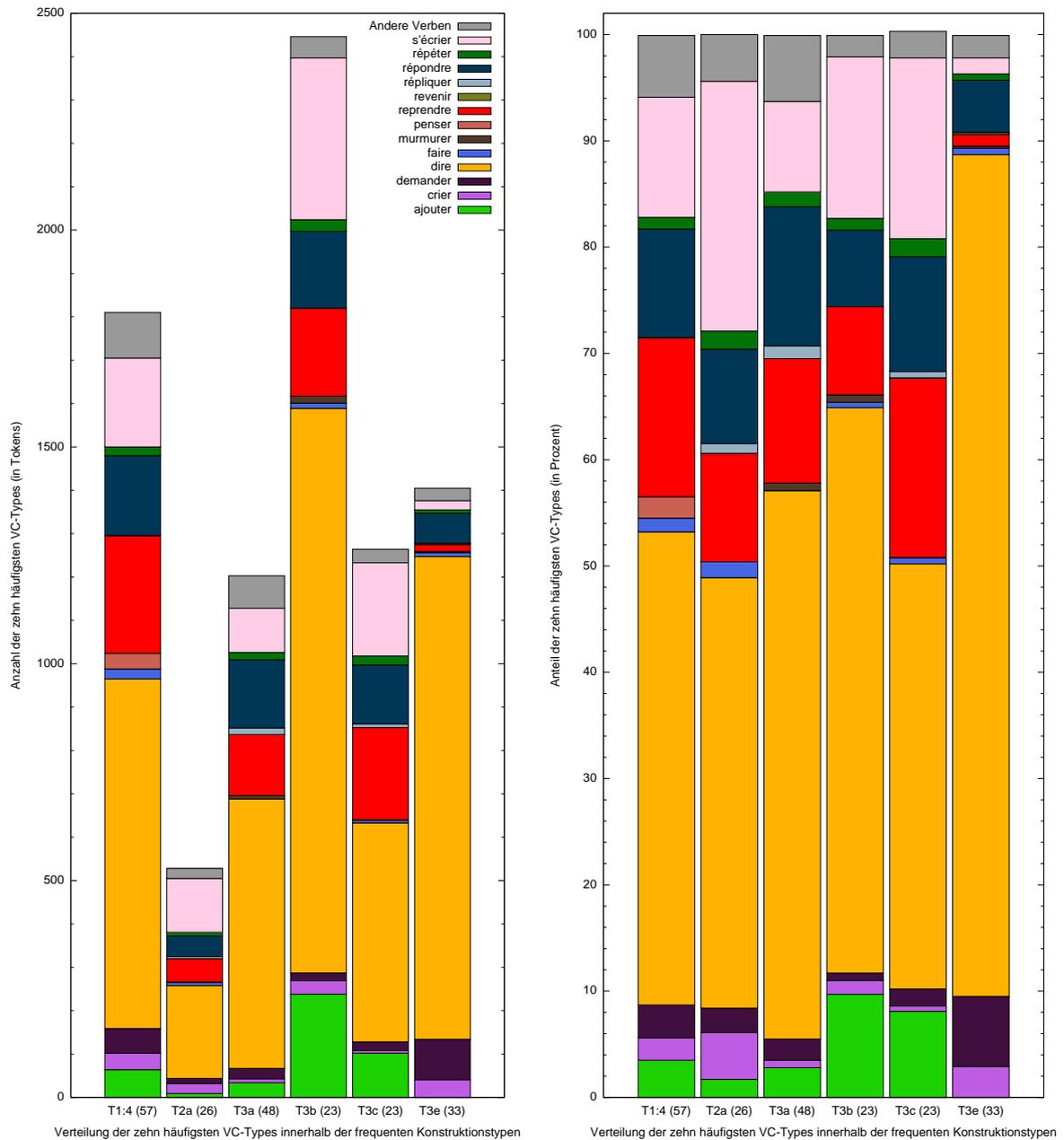


Abbildung 4.23: Verteilung der zehn häufigsten VC-Types pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Victor Hugo**, **Stendhal**, **George Sand** und **Eugène Sue**. Die senkrechte Achse gibt die Anzahl der VC-Tokens wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Types an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

(5350) – *D'où venez-vous comme ça ? dit le portier en sortant à demi de sa loge, tenant à la main un fusil à deux coups qu'il s'occupait de charger, et en examinant l'ouvrière d'un regard soupçonneux.*

(5351) – *Je viens de parler à Mme la supérieure, répondit timidement la Mayeux.*¹⁷⁹

In anderen Fällen geht der Redebegleitung mit *répondre* ein Turn voraus, dessen *verbum communicandi* keine Frage erwarten lässt. Dabei handelt es sich meistens um das semantisch neutrale *dire*:

(5098) – *La condition d'aller en journée est malheureusement indispensable, dit Florine.*

(5099) – *Alors, je dois renoncer à cet espoir, répondit timidement la Mayeux.* . . .¹⁸⁰

Vereinzelt finden sich Beispiele, in denen weder ein Interpunktionszeichen noch ein *verbum communicandi* eine Frage anzeigt, sondern nur eine indirekte Fragehandlung erfolgt:

(2929) – *Eh bien, ce soir arrivera, Dieu merci ! et peut-être l'attendez-vous avec autant d'impatience que moi. Peut-être, ce soir, Mme Bonacieux visitera-t-elle le domicile conjugal.*

(2930) – *Mme Bonacieux n'est pas libre ce soir, répondit gravement le mari ; elle est retenue au Louvre par son service.*¹⁸¹

Anders sieht der Gebrauch der beiden Verben innerhalb der Dativkonstruktion (T3e) aus. Das Verhältnis von *répondre* und *demander* ist in den Gruppen Hugo/Stendhal/Sand/Sue und Honoré de Balzac¹⁸² weitgehend ausgeglichen, während es sich bei Dumas/Verne/de Goncourt/Flaubert, Zola sowie Prévost/Alain-Fournier/de Goncourt/Flaubert ins Gegenteil verkehrt.¹⁸³ Lediglich bei Voltaire und de Saint-Pierre bleibt die Überzahl von *répondre* erhalten.

Für die Dativkonstruktion (T3e) ergeben sich bezüglich der Verteilung der Verben weitere Besonderheiten. *Dire* wird innerhalb von T3e so häufig wie in keinem der übrigen Konstruktionstypen verwendet, und zwar in allen Autorengruppen bis auf Émile Zola zu ähnlich großen Anteilen¹⁸⁴. Dies ist unter anderem der Tatsache geschuldet, dass ein

¹⁷⁹ Eugène Sue, *Le juif errant* (1845).

¹⁸⁰ Eugène Sue, *Le juif errant* (1845).

¹⁸¹ Alexandre Dumas, *Les trois mousquetaires* (1844).

¹⁸² Hugo/Stendhal/Sand/Sue: *demander* 6,6% vs. *répondre* 4,9%, Balzac: *demander* 8,9% vs. *répondre* 5,5%.

¹⁸³ Hugo/Stendhal/Sand/Sue: *demander* 21,4% vs. *répondre* 3,7%, Zola: *demander* 21,1% vs. *répondre* 2%, Prévost/Alain-Fournier/de Goncourt/Colette: *demander* 15,6% vs. *répondre* 2,3%.

¹⁸⁴ Bei Émile Zola beläuft sich der Anteil von *dire* innerhalb T3e auf 49,8%. In den übrigen Autorengruppen beträgt er zwischen 65,8% (Gruppe Prévost/Alain-Fournier/Proust/Colette) und 80% (Voltaire und de Saint-Pierre). Der Anteil ist in den Romanen der zweiten Korpushälfte grundsätzlich niedriger als in den frühen Texten.

verbum communicandi zur Verwendung innerhalb der Dativkonstruktion, mittels derer auf einen Gesprächspartner Bezug genommen wird, die entsprechende Valenz aufweisen muss. Neben *demander*, *répondre* und *crier* erfüllt *dire* dieses Kriterium perfekt, da es anders als die ersten drei semantisch neutral ist und dadurch sehr leicht mit weiteren Konstruktionstypen kombiniert werden kann. Umgekehrt sind die Redeverben *s'écrier* und *reprendre*, letzteres im Sinne von «[p]rendre la parole après un silence, pour dire quelque chose»¹⁸⁵ bzw. «se remettre à dire quelque chose après qu'on s'était arrêté de parler»¹⁸⁶, selten oder gar nicht unter den zehn häufigsten Verben zu finden, weil aufgrund ihrer Intransitivität kein Dativobjekt möglich ist. Dass trotzdem Konstruktionen mit *s'écrier* oder *reprendre* gefunden und der Dativkonstruktion zugeordnet werden (siehe die folgenden zwei Beispiele), liegt daran, dass das Programm nicht erkennen kann, dass die Fundstelle ungültig ist:

(3908) – *J'ai bien du malheur ! s'écria Orugix à ce récit.*¹⁸⁷

(2036) – *Je ne sais pas seulement votre nom, reprit l'homme au manteau avec sa voix de sépulcre. Vous avez un rendez-vous ce soir.*¹⁸⁸

Obwohl *s'écrier* per Definition¹⁸⁹ intransitiv ist, stößt man hin und wieder auch auf transitiven Gebrauch:

(4780) – *Il est donc vrai ! s'écria Mme Dubreuil à la laitière, qui tenait toujours Fleur-de-Marie par le bras, vous osez parler de la sorte à l'amie de ma fille !*¹⁹⁰

Auch für *reprendre* gibt es Beispiele mit transitivem Gebrauch:

(4455) – *Voulez-vous venir, mes chères demoiselles ? reprit Mme Grivois aux deux jeunes filles, qui avaient à la dérobee fait quelques caresses à Rabat-Joie, tant elles étaient touchées de son instinct ; on va vous conduire auprès de Mme votre parente, et je reviendrai vous prendre dans une demi-heure : cocher, retenez bien le chien.*¹⁹¹

Die Beispiele (4780) und (4455) sind jedoch die absolute Ausnahme. Unabhängig von der Valenz des verwendeten *verbum communicandi* wäre ein einfaches Ausschlusskriterium für Beispiele wie (3908) oder (2036) die Belebtheit bzw. Unbelebtheit des indirekten

¹⁸⁵ Rey-Debove & Rey (1993), S. 1942.

¹⁸⁶ Quelle: TLFi (<http://atilf.atilf.fr/>, Eintrag unter II A.2).

¹⁸⁷ Victor Hugo, *Han d'Islande* (1823).

¹⁸⁸ Victor Hugo, *Notre-Dame de Paris* (1831).

¹⁸⁹ «Dire d'une voix forte et émue (Rey-Debove & Rey (1993), S. 715.) bzw. «Prononcer quelques paroles d'une voix forte en manifestant une émotion, un sentiment spontané, etc.»(Quelle: TLFi (<http://atilf.atilf.fr/>, Eintrag unter B.1)).

¹⁹⁰ Eugène Sue, *Les mystères de Paris* (1842).

¹⁹¹ Eugène Sue, *Le juif errant* (1845).

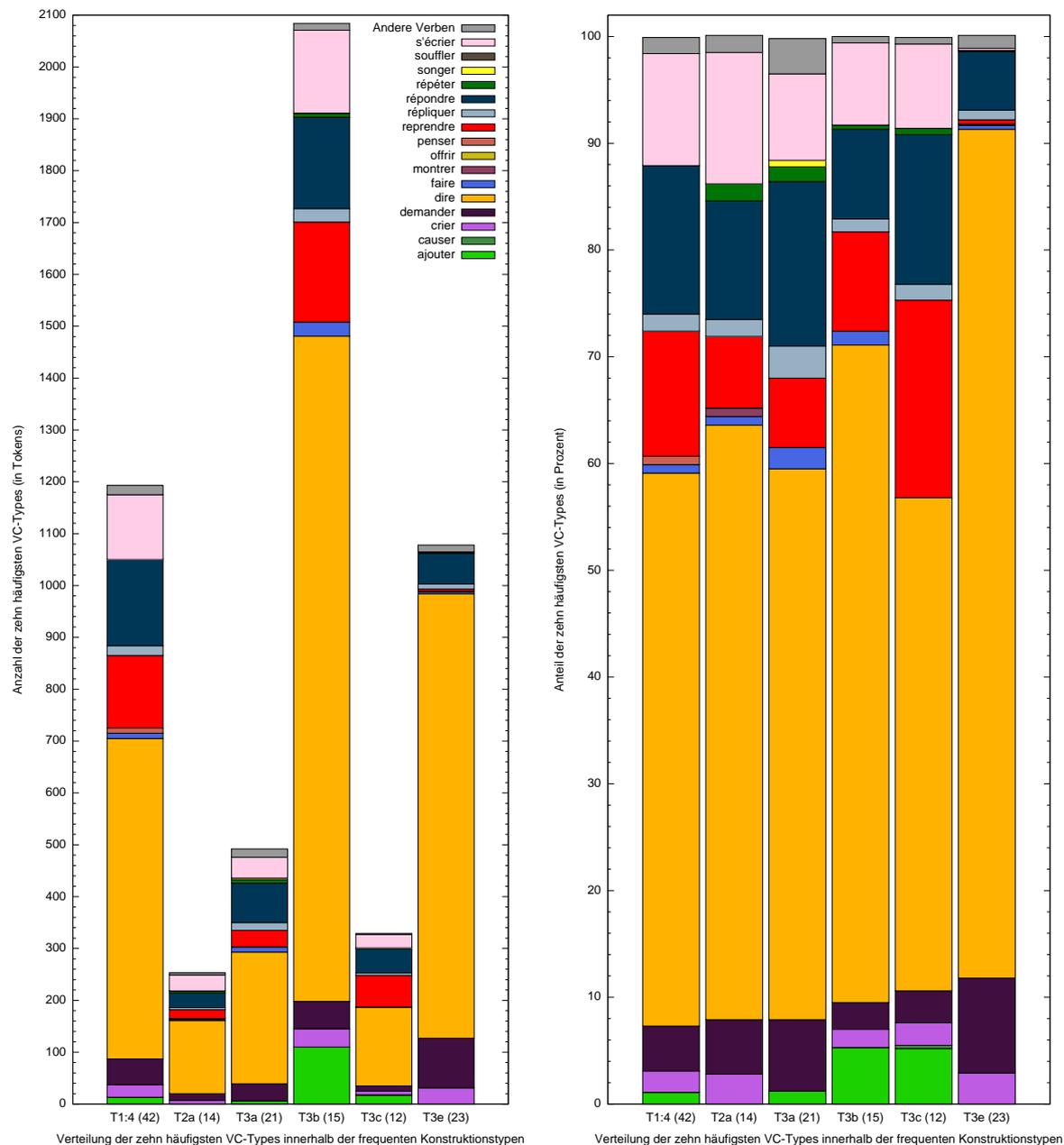


Abbildung 4.24: Verteilung der häufigsten VC-Types pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Honoré de Balzac**. Die senkrechte Achse gibt die Anzahl der VC-Tokens wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Types an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

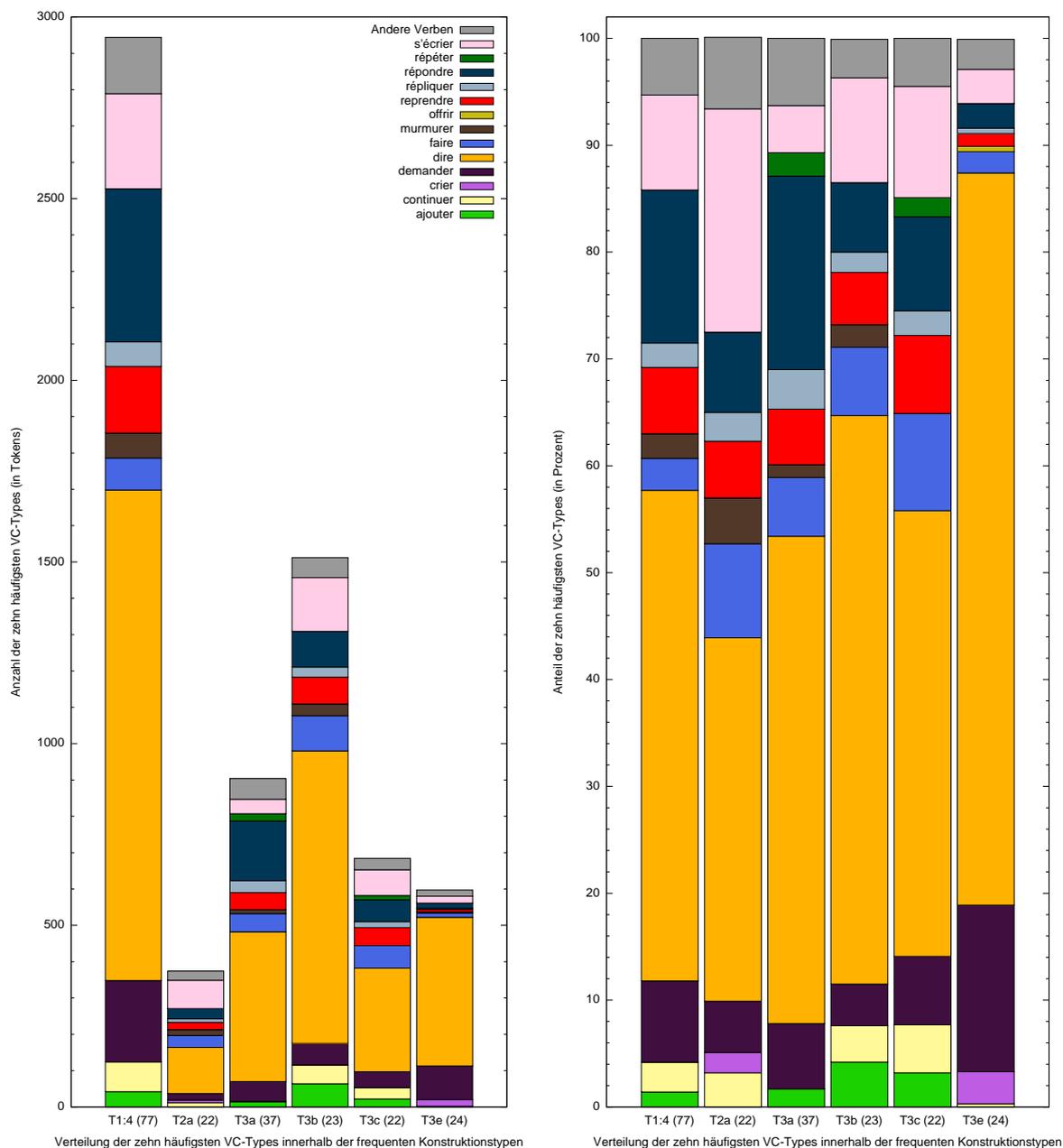


Abbildung 4.25: Verteilung der häufigsten VC-Typen pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Alexandre Dumas**, **Jules Verne**, **Jules und Édmond de Goncourt** und **Gustave Flaubert**. Die senkrechte Achse gibt die Anzahl der VC-Tokens wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Typen an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

Objekts, d.h. der Substantive *récit* und *manteau*. Da diese jedoch durch TreeTagger nicht bestimmt werden kann, müssen derlei Fehltreffer in Kauf genommen werden.

Aufschlussreich ist auch ein Blick auf den zuoberst liegenden grau gefärbten Stapel, der all diejenigen VC-Tokens enthält, die zu keinem der zehn häufigsten VC-Typen gehören. Was auf das Verhältnis der VC-Tokens zu der tatsächlichen Anzahl von VC-Typen zutrifft, gilt auch für die Rest-Tokens. Deren Anzahl sagt kaum etwas über deren Zusammensetzung aus: So können sich entweder viele VC-Tokens auf wenige Typen verteilen, oder es sind umgekehrt viele VC-Typen mit nur wenigen VC-Tokens enthalten. Dies variiert zudem von Autorengruppe zu Autorengruppe und von Konstruktionstyp zu Konstruktionstyp. Während die Gruppe Voltaire und Bernardin de Saint-Pierre in keinem Konstruktionstyp mehr als sechs VC-Typen und folglich keine Rest-Tokens aufweist (vgl. Abbildung 4.22), liegt der Anteil für die übrigen Autorengruppen zwischen 0,4% und 14,1%¹⁹². Für die Romane Honoré de Balzacs wurden insgesamt die wenigsten Rest-Tokens ermittelt. Für T1 sind dies nur 18 VC-Tokens. Die Anzahl der Rest-Tokens nimmt jedoch allmählich, d.h. zum Ende des Untersuchungszeitraums hin, zu. Im Vergleich zu Honoré de Balzac hat Émile Zola für T1 – beide weisen ähnlich viele VC-Tokens (1193 für Balzac bzw. 1119 für Zola) für diesen Konstruktionstyp auf – mit 137 mehr als siebenmal soviel Rest-Tokens. Für die Gruppe Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette verhält es sich ähnlich: Von insgesamt deutlich weniger VC-Tokens in T1 (263) entfallen 47 auf den Rest.

Geht man nun eine Ebene tiefer und vergleicht die Schaubilder der einzelnen Autorengruppen, so ergibt sich ein differenzierteres Bild.

Die Sonderstellung der Autorengruppe **Voltaire** und **Bernardin de Saint-Pierre** fällt bei Betrachtung von Abbildung 4.22 sofort ins Auge. Die Vereinigungsmenge umfasst nur elf Verben, von denen pro Konstruktionstyp nur maximal sechs vorkommen. Die Redebegleitung wird fast ausschließlich mittels des Basistyps (T1) und der Dativkonstruktion (T3e) konstruiert, was anhand der linken Graphik von Abbildung 4.22 deutlich wird. Konstruktionen wie die Adverbial- (T3a) oder die Dativkonstruktion (T3e), die bei allen anderen Autoren sehr populär sind, kommen nur selten vor. Auch die Variabilität der *verba communicandi*¹⁹³ ist aufgrund weniger VC-Tokens, die sich auf wenige VC-Typen verteilen, gering. Verwendet werden vor allem die Verben *dire*, *répondre* und *répliquer*. In T1 sind außerdem *s'écrier* und *repandre* häufig. Letzteres kann jedoch aufgrund seiner Valenz nicht in einem redebegleitenden *discours citant* mit Dativkonstruktion (T3e)

¹⁹² Die Anzahl der Rest-Tokens wird in Beziehung gesetzt zu der Anzahl der Tokens für den jeweiligen Konstruktionstyp.

¹⁹³ Zur Variabilität der *verba communicandi* siehe auch Kapitel 4.2.4.

stehen. Das Verb *s'écrier* ist in diesen frühen Texten das einzige, das innerhalb der Redebegleitung die Stimme der sprechenden Figur hinsichtlich Lautstärke und Ausdruck von Emotionen moduliert. Die in den anderen Autorengruppen teilweise sehr häufigen Verben *crier* und *ajouter* sind gar nicht enthalten, während *demander* und *continuer* nur in je einem Konstruktionstyp, d.h. T1 bzw. T3e, vorkommen.

In der Gruppe **Victor Hugo, Stendhal, George Sand** und **Eugène Sue** sind nicht nur mehr *verba communicandi* enthalten, sondern auch die Ausgestaltung des redegleitenden discours citant ist durch den häufigen Einsatz weiterer Konstruktionstypen abwechslungsreicher geworden. Fast alle Verben der Vereinigungsmenge kommen mit Ausnahme der Dativkonstruktion (T3e) zu ungefähr gleichen Teilen in den Konstruktionen vor, allen voran *dire*. Sehr häufig sind außerdem *s'écrier*, *reprendre* und *répondre*. *Ajouter*, *répéter*, *crier* und *murmurer* gibt es zwar in der Gruppe Voltaire und de Saint-Pierre noch nicht, doch hier gehören sie erstmals zu den zehn häufigsten Verben. Innerhalb von T1 gehört dazu auch *penser*¹⁹⁴:

(1523) – *Tout cela est fort inquiétant, pensait-il.*¹⁹⁵

Zu den expressiven Verben gehören *s'écrier*, *crier* und *murmurer*. Von diesen ist *s'écrier* gerade in dieser Autorengruppe mit Anteilen zwischen 7,7% für T3b (*gérondif*-Konstruktion) und 12,3% für T2a (Konstruktion mit Adjektiv oder *participe passé*) sehr stark vertreten. *Murmurer* dagegen befindet sich nur in der Adverbialkonstruktion (T3a) und der *gérondif*-Konstruktion mit geringen Anteilen zwischen 0,5% und 0,7% unter den zehn häufigsten Redeverben. Ein weiteres Verb, das sich allmählich zu einem gängigen Redeverb entwickelt, ist *faire*¹⁹⁶. Bis auf T3a (Adverbialkonstruktion) ist es in allen Konstruktionstypen enthalten, allerdings mit nur geringen Anteilen zwischen 0,5% (T3b) und 1,5% (T2a). Innerhalb der Dativkonstruktion (T3e) dominiert *dire* mit einem Anteil von 79,2%, gefolgt von *demander* mit 6,6% und *répondre* mit 4,9%.

In den 16 untersuchten Romanen von **Honoré de Balzac** sieht die Verteilung der häufigsten zehn Verben innerhalb der frequenten Konstruktionstypen mit kleinen Abweichungen ähnlich aus wie in der Gruppe zuvor. Die Vereinigungsmenge der Verben für alle Konstruktionstypen umfasst 16 Verben, von denen *dire* mit Abstand das häufigste in allen Konstruktionstypen ist. Zu ungefähr gleichen Teilen folgen *reprendre*, *répondre* und *s'écrier*. *Ajouter* und *répéter* sind nicht mehr so häufig vertreten wie noch in der zuvor

¹⁹⁴ *Penser* kommt in den anderen Konstruktionstypen zwar auch vor, gehört dort aber nicht zu den zehn häufigsten Verben.

¹⁹⁵ Victor Hugo, *Han d'Islande* (1823).

¹⁹⁶ Während die Verbreitung von *faire* als Redeverb ab der zweiten Korpushälfte generell zunimmt, ist es erstaunlicherweise bei Émile Zola nicht unter den zehn häufigsten *verba communicandi* zu finden (vgl. Abbildung 4.26)!

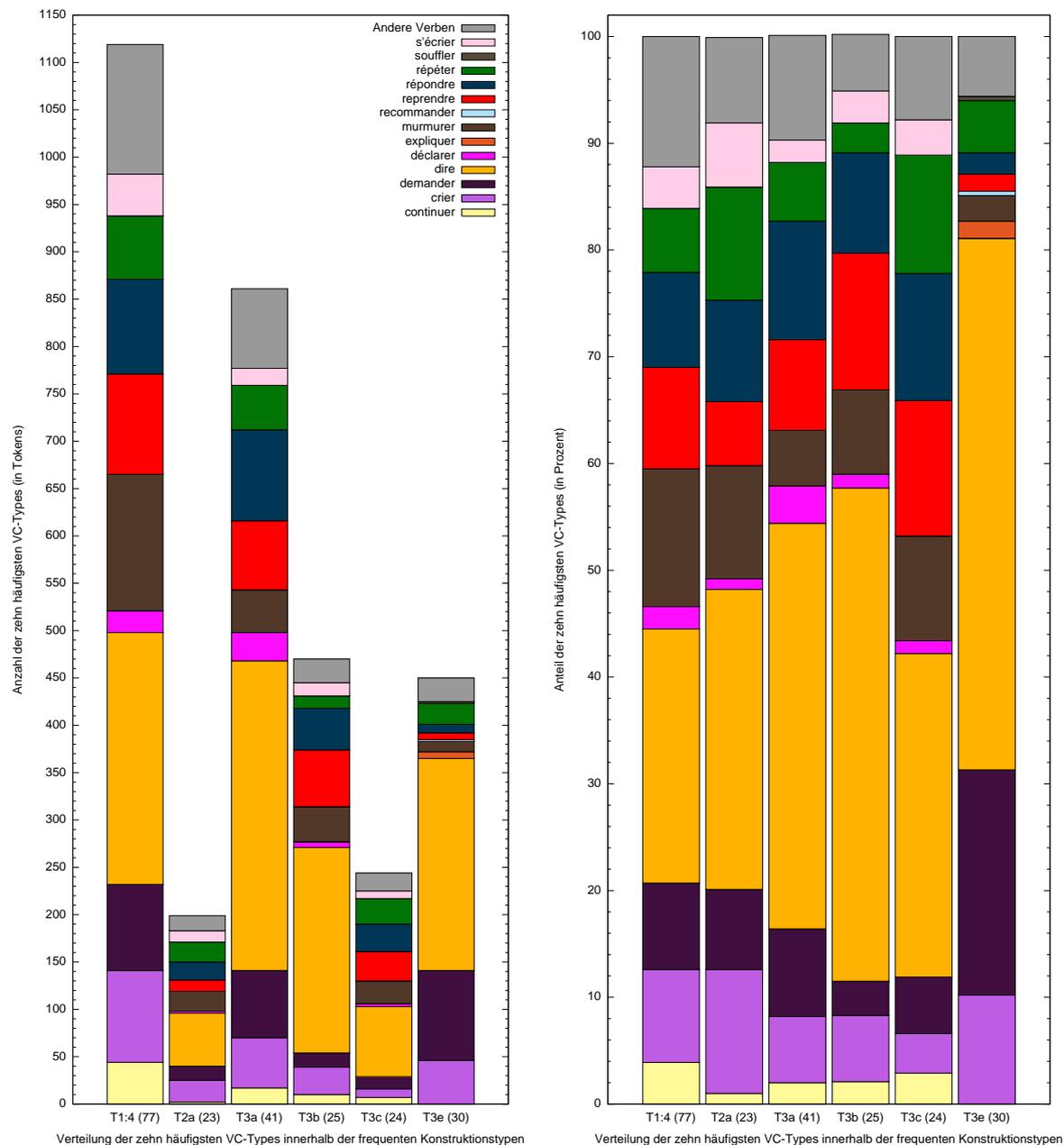


Abbildung 4.26: Verteilung der zehn häufigsten VC-Typen pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Émile Zola**. Die senkrechte Achse gibt die Anzahl der VC-Tokens wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Typen an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

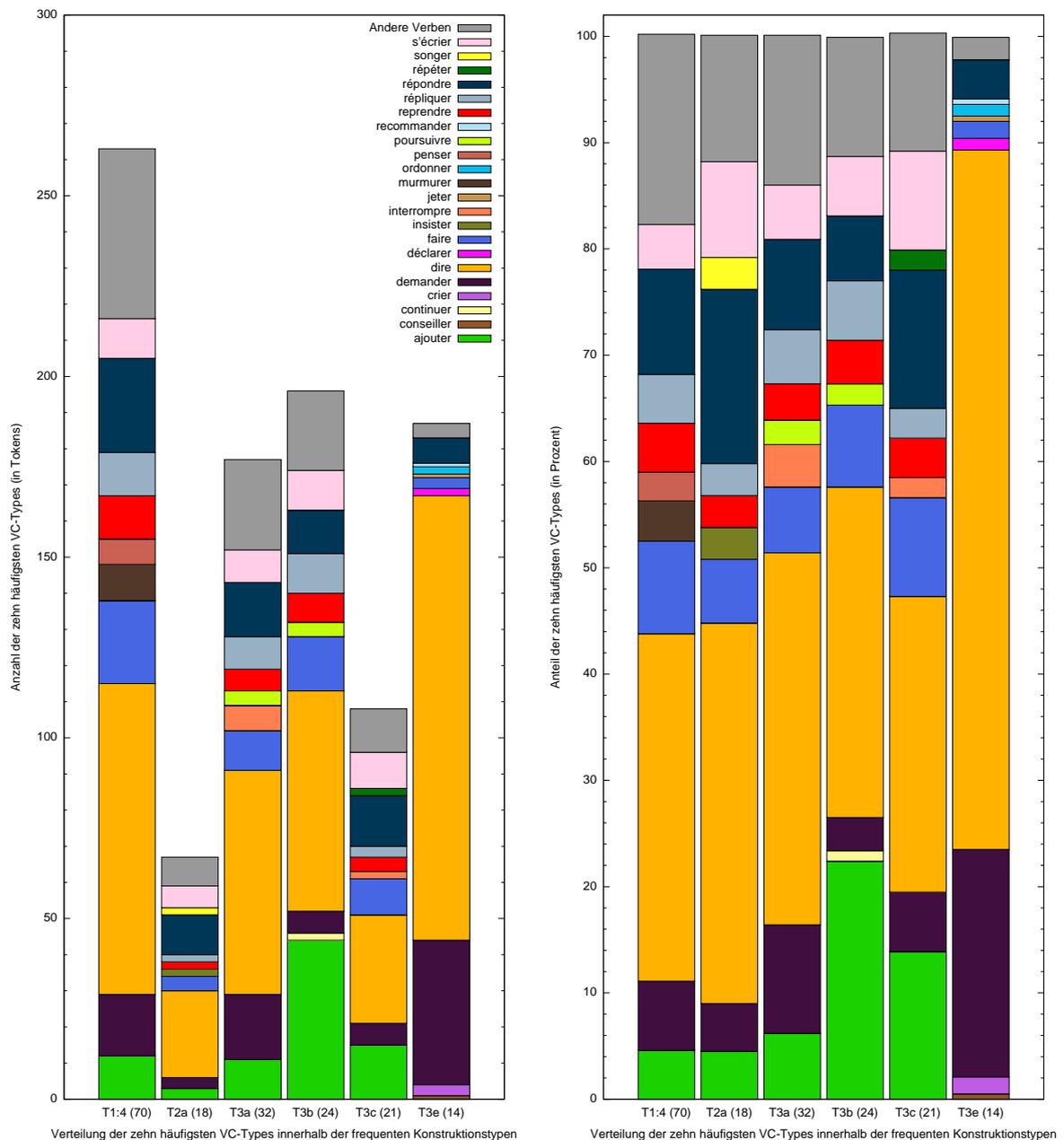


Abbildung 4.27: Verteilung der zehn häufigsten VC-Typen pro Konstruktionstyp (Auflösung der Kürzel mittels Übersicht auf S. 113) bei **Marcel Prévost**, **Henri Alain-Fournier**, **Marcel Proust** und **Colette**. Die senkrechte Achse gibt die Anzahl der VC-Tokens wieder, während entlang der horizontalen Achse die einzelnen Konstruktionstypen dargestellt werden. Die Zahl in Klammern hinter jedem Konstruktionstyp gibt die Anzahl der VC-Typen an. Zwecks besserer Darstellung wurden die Werte für T1 jeweils geviertelt.

besprochenen Autorengruppe, dafür hat *demander* größere Anteile, und auch *répliquer* gehört in allen Konstruktionstypen zu den zehn häufigsten *verba communicandi*. Auch *faire* verzeichnet einen geringen Zuwachs und ist in allen Konstruktionstypen mit Werten zwischen 0,4% (T3e) und 2% (T3a) vertreten. Im Gegensatz zur vorherigen Gruppe befindet sich *continuer* nicht unter den zehn häufigsten Verben. Bei Balzac gibt es nur zwei Verben, die die Stimmqualität näher beschreiben: *s'écrier* und *crier*. Zwar wurde auch *souffler* ermittelt, doch fällt dieses mit einem Anteil von nur 0,1% (T3e) praktisch nicht ins Gewicht. Die Dativkonstruktion ist auch für diesen Autor durch die Dominanz des Verbs *dire* (79,5%) gekennzeichnet, das vor allem mit den Verben *demander*, *répondre* und *crier* variiert wird.

In der Autorengruppe **Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt** und **Gustave Flaubert** umfasst die Vereinigungsmenge zwar nur 13 Verben, jedoch wirkt in Abbildung 4.25 die rechte Graphik mit der prozentualen Darstellung insgesamt „bunter“. Das kommt dadurch zustande, dass der Anteil von *dire* und auch anderen Verben in fast allen Konstruktionstypen zurückgeht und dafür andere *verba communicandi* hinzukommen. So reduzieren sich beispielsweise die Vorkommen von *reprendre*, die nur noch einen Anteil zwischen 4,9% (T3b) und 7,3% (T3c) ausmachen.¹⁹⁷ Ähnlich verhält es sich auch mit *ajouter*. Im Gegenzug liegt der Anteil von *continuer*, *demander*, aber auch *répliquer* und *murmurer* deutlich höher als in allen bisher betrachteten Autorengruppen. *Faire* legt ebenfalls zu und ist mit Anteilen zwischen 2% (T3e) und 9,1% (T3c) vertreten. Während *s'écrier* nach wie vor in allen Konstruktionstypen zu den vier häufigsten Verben gehört, ist *crier*, bedingt durch die Zunahme anderer Verben, nur noch innerhalb der Konstruktion mit Adjektiv oder *participe passé* (T2a) bzw. der Dativkonstruktion (T3e) Teil der zehn häufigsten Redeverben. Mit *murmurer* ist außerdem ein drittes expressives *verbum communicandi* enthalten.

Bei den untersuchten Romanen von **Émile Zola** (vgl. Abbildung 4.26) fällt auf, dass für alle Konstruktionstypen außer der Dativkonstruktion (T3e) dieselben zehn Verben¹⁹⁸ am häufigsten sind, auch wenn sich die Menge der VC-Tokens von Typ zu Typ mehr oder weniger deutlich unterscheidet. An die Stelle des Konstruktionstyps T3b (*gérondif*-Konstruktion), der in allen Autorengruppen nach T1 (außer bei Voltaire und Bernardin de Saint-Pierre) die häufigste Konstruktion bildet, tritt außerdem die Adverbialkonstruktion (T3a), die ungefähr doppelt so viele VC-Tokens enthält wie T3b. Wie schon bei Dumas/

¹⁹⁷ In der Gruppe Victor Hugo, Stendhal, George Sand und Eugène Sue beträgt der Anteil von *reprendre* noch zwischen 8,3% und 16,9%.

¹⁹⁸ Diese sind *continuer*, *crier*, *demander*, *dire*, *déclarer*, *murmurer*, *reprendre*, *répondre*, *répéter* und *s'écrier*.

Verne/de Goncourt/Flaubert fällt der Anteil von *dire* insgesamt geringer aus, was dazu führt, dass hier mehr VC-Tokens in ähnlich großer Zahl enthalten sind als in anderen Gruppen, so z.B. *crier*, *demander*, *murmurer*, *reprendre*, *répondre* und *répéter*. Während weder *ajouter*, noch *répliquer* oder *faire* zu den zehn häufigsten *verba communicandi* gehören, ist der Anteil von *répéter* deutlich größer als bei allen anderen Autorengruppen, und auch *continuer* zählt, anders als in den meisten anderen Gruppen, zu den meistverwendeten. Bei Zola gehört in allen Konstruktionstypen außer T3e erstmalig *déclarer* mit Anteilen bis zu 3,5% (T3a) zu den häufigsten *verba communicandi*.¹⁹⁹ Interessant ist bei Zola außerdem, dass sich das Verhältnis von *s'écrier* und *crier*, das in den zuvor betrachteten Autorengruppen zugunsten von *s'écrier* ausfiel, umkehrt. Somit ist *crier* mit maximal 11,6% (T2a) so stark wie in keiner anderen Gruppe.

Mit einer Vereinigungsmenge von insgesamt 22 Verben liegt die Gruppe der Autoren **Marcel Prévost**, **Henri Alain-Fournier**, **Marcel Proust** und **Colette** (s. Abb. 4.27) deutlich über dem Schnitt von 14 Verben. Hieraus ergibt sich eine – auch durch die Farben im Schaubild auffallende – heterogene Zusammensetzung der in den Konstruktionstypen verwendeten Redeverben. Dieser Effekt wird dadurch verstärkt, dass *dire* in allen Konstruktionstypen bis auf T3e (Dativkonstruktion) mit Anteilen zwischen 27,8% (T3c) und 35,8% (T2a) vergleichsweise selten auftritt. In allen sechs Konstruktionstypen vertreten sind *dire*, *demander*, *faire* und *répondre*. In allen Konstruktionstypen außer T3e (Dativkonstruktion) wurden außerdem *ajouter*, *s'écrier*, *répliquer* und *reprendre* ermittelt. Wie bei Dumas/Verne/de Goncourt/Flaubert wird *reprendre* seltener gebraucht, dafür gehört *répliquer* wieder zu den zehn häufigsten *verba communicandi*. Als expressives Verb wird vor allem *s'écrier* eingesetzt, während *crier* nur innerhalb von T3e und *murmurer* nur in T1 zu den zehn häufigsten Verben gehört. Einzelne Verben sind am Ende des Untersuchungszeitraum erstmals so beliebt, dass sie zu den häufigsten Redeverben zählen. Allerdings treten sie nur innerhalb einzelner Konstruktionstypen gehäuft auf. So wird bei den Autoren in T1 öfter *penser* eingesetzt, in T2a hingegen *insister* und *songer*. Neu sind in T3a und T3b außerdem *poursuivre*, genauso wie *interrompre* in T3a und T3c.

¹⁹⁹ Danach ist es nur noch für T3e in der Gruppe Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette (s. Abb. 4.27, Konstruktionstyp T3e) hochfrequent.

4.2.4 Type-Token-Relation der *verba communicandi*

Für die Bewertung von Textkorpora hinsichtlich ihres Wortschatzes ist die Type-Token-Relation (TTR) von Bedeutung. Diese errechnet sich, indem die Anzahl der Types, d.h. die den Wortformen zugrundeliegenden Lexeme, durch die Anzahl der Tokens, d.h. die einzelnen Wortformen, dividiert und der Quotient mit 100 multipliziert wird. Die Angabe des Wertes erfolgt demnach in Prozent und beträgt höchstens 100%. Je länger ein Text, desto niedriger ist seine TTR, „weil die Wiederholungen schneller als die neuen Wortformen zunehmen“²⁰⁰.

Das Prinzip der Type-Token-Relation kann auch auf die Beziehung zwischen VC-Types und VC-Tokens innerhalb der Redebegleitung angewendet werden. Auf diese Weise ist feststellbar, welche Autoren sich um einen besonders variantenreichen Einsatz von *verba communicandi* bemühen.

Eine Gegenüberstellung aller im Arbeitskorpus ermittelten VC-Tokens und VC-Types innerhalb der Redebegleitung, unabhängig davon, welcher Konstruktionstyp jeweils vorliegt, erfolgt in Abbildung 4.28. Die TTR der *verba communicandi* und damit deren Variabilität ist per Definition umso größer, je enger die Anzahl der VC-Tokens und die Anzahl der VC-Types beieinanderliegen. Die TTR ist umgekehrt umso niedriger, je mehr VC-Tokens insgesamt enthalten sind. Die Anzahl der VC-Tokens kann zwar beliebig zunehmen, die Zahl der möglichen VC-Types jedoch nicht proportional dazu anwachsen. Besonders deutlich wird dieser Zusammenhang anhand der Texte Nr. 20²⁰¹, 25²⁰², 27²⁰³, 29²⁰⁴ und 33²⁰⁵ (vgl. Abbildung 4.28). Obwohl diese zwar im Vergleich zu den übr-

²⁰⁰ Messelken (1996), S. 211. Um diesen Effekt bei sehr langen Texten etwas abzuschwächen und eine bessere Vergleichbarkeit von kürzeren mit längeren Texten zu erhalten, gibt es zusätzlich die Möglichkeit zur Berechnung einer standardisierten TTR. Dazu wird ein Text von Anfang bis Ende in Abschnitte von je 1000 Tokens – bei kurzen Texten entsprechend weniger – aufgeteilt. Für jeden Abschnitt wird die TTR separat ermittelt und aus diesen allen ein Mittelwert berechnet (mit einer solchen standardisierten TTR arbeitet z.B. auch die Software *Wordsmith Tools*). Grundsätzlich ließe sich eine standardisierte TTR auch bei der Beurteilung des Verhältnisses von VC-Types zu VC-Tokens anwenden. Denkbar sind zwei Methoden: Bei Methode A würde der gesamte Text in Subtexte mit einer Länge von 1000 Tokens aufgeteilt und für jeden Abschnitt die enthaltenen Dialogstellen ermittelt. Anhand der darin enthaltenen VC-Tokens ließe sich dann auf die VC-Types schließen und die TTR für jeden Subtext berechnen. Die standardisierte TTR würde man schließlich aus der Mittelung aller TTRs erhalten. Bei Methode B würde man die VC-Tokens in der Reihenfolge ihres Auftretens sammeln, diese Liste in Teillisten von 10 bis 100 VC-Tokens unterteilen, für jede Teilliste die VC-Types ermitteln und anschließend die TTR berechnen. Aus den TTRs für alle Teillisten wird auch hier wieder der Mittelwert errechnet. Beide Methoden bringen keinen erkennbaren Vorteil für die Auswertung, weshalb auf eine Standardisierung verzichtet wird.

²⁰¹ Eugène Sue, *Les mystères de Paris* (1842).

²⁰² Alexandre Dumas, *Les trois mousquetaires* (1844).

²⁰³ Eugène Sue, *Le juif errant* (1845).

²⁰⁴ Alexandre Dumas, *Le comte de Monte-Christo* (1846).

²⁰⁵ Alexandre Dumas, *Le collier de la reine* (1849).

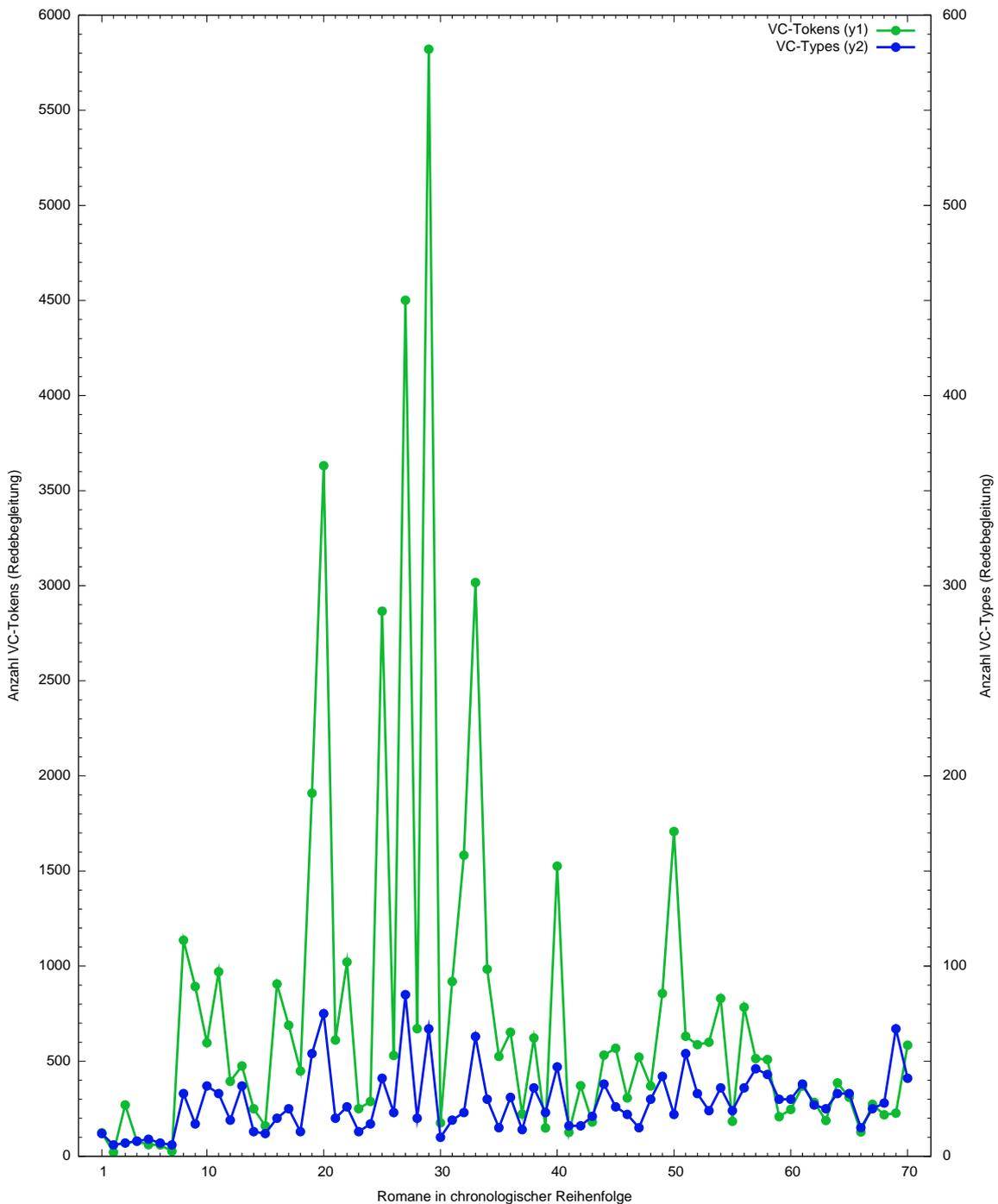


Abbildung 4.28: Gegenüberstellung von Tokens und Types innerhalb der Redebegeleitung für das gesamte Korpus und in absoluten Zahlen (unabhängig vom Konstruktionstyp). Die grüne Kurve bildet die Anzahl der VC-Tokens für jeden Text ab. Die Werte sind an der linken senkrechte Achse abzulesen. Die blaue Kurve zeigt an, zu wie vielen Types die Tokens zusammengefasst werden. Da deutlich weniger Types als Tokens vorkommen, bedarf es zwecks besserer Darstellung im Schaubild einer zweiten senkrechten Achse (siehe rechter Bildrand), mit deren Hilfe die Anzahl der Types bestimmt werden kann. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

gen Romanen innerhalb des Korpus überdurchschnittlich viele VC-Tokens und auch VC-Types aufweisen, bewegt sich die TTR der *verba communicandi* dennoch nur zwischen 1,2% und 2,1%. Umgekehrt liegt eine hohe TTR für diejenigen Romane vor, in denen wie in Abbildung 4.28 die grüne Kurve (VC-Tokens) sehr nahe an die blaue Kurve (VC-Types) heranreicht, was vor allem bei Texten gegen Ende des Untersuchungszeitraums zwischen 1885 und 1922, d.h. bei den Romanen Nr. 57-70, der Fall ist.

Während Abbildung 4.28 die tatsächlichen Vorkommen der VC-Tokens und VC-Types für die gesamte Redebegleitung zeigt, wird in Abbildung 4.29 die TTR dieser beiden Werte dargestellt, was den Vergleich der Romane in Bezug auf die Variabilität der *verba communicandi* ermöglicht.

Dass die meisten Texte von Voltaire (Nr. 1-6) für die Redebegleitung eine TTR von knapp 10% und mehr aufweisen, ist nicht weiter verwunderlich, denn in diesen Texten gibt es nur wenig direkte Rede, so dass sich wenige VC-Tokens auf wenige VC-Types verteilen. Während die TTR für *Paul et Virginie* (1787) von Bernardin de Saint-Pierre (Text Nr. 8) noch bei 22,3% liegt, fällt sie danach für *Han d'Islande* (1823) von Victor Hugo auf 2,9% und übersteigt erstmals knapp vier Jahrzehnte später in *Salammô* (1862) von Gustave Flaubert (Text Nr. 39) wieder die 10-Prozent-Marke. Während dieser knapp 40 Jahre enthalten die Texte Nr. 8-38 meistens viele Dialogpartien mit entsprechend vielen VC-Tokens und – so belegen es die absoluten Zahlen – vielen VC-Types. Weil die VC-Tokens jedoch um ein Vielfaches zahlreicher sind als die VC-Types und diese nicht beliebig zunehmen können, liegt die TTR in nur wenigen Romanen, d.h. in neun von 31 Texten, über 5%²⁰⁶.

Zwischen 1862 und 1887 (Texte Nr. 39-58) folgen die Texte bezüglich der TTR keinem eindeutigen Trend, sondern es liegt ein stetiges Auf und Ab vor. Allerdings bewegen sich die Maxima auf einem höheren Niveau als noch für den davorliegenden Zeitraum, und auch die Minima liegen mit nur wenigen Ausnahmen höher als noch zuvor. Mit einer erhöhten TTR zwischen 11,7% und 15,4% fallen vor allem die Texte Nr. 39, *Salammô* (1862) von Gustave Flaubert, Nr. 41, *Germinie Lacerteux* (1865) von Jules und Édmond de Goncourt, Nr. 43, *Thérèse Raquin* (1867) von Émile Zola sowie Nr. 55, *Bouvard et Pécuchet* (1881) von Gustave Flaubert auf²⁰⁷.

²⁰⁶ Dies trifft zu auf die Texte Nr. 10: *La peau de chagrin* (1831), Nr. 13: *Le médecin de campagne* (1833), Nr. 14: *La duchesse de Langeais* (1835), Nr. 15: *La fille aux yeux d'or* (1835), Nr. 23: *Le colonel Chabert* (1844), Nr. 24: *Le lys dans la vallée* (1844) – alle von Honoré de Balzac –, Nr. 30: *La mare au diable* (1846) und Nr. 37: *François le Champi* (1850) – beide von George Sand – und Nr. 38, *Le médecin de campagne* (1857) von Gustave Flaubert.

²⁰⁷ Vgl. Tabelle B.1 im Anhang.

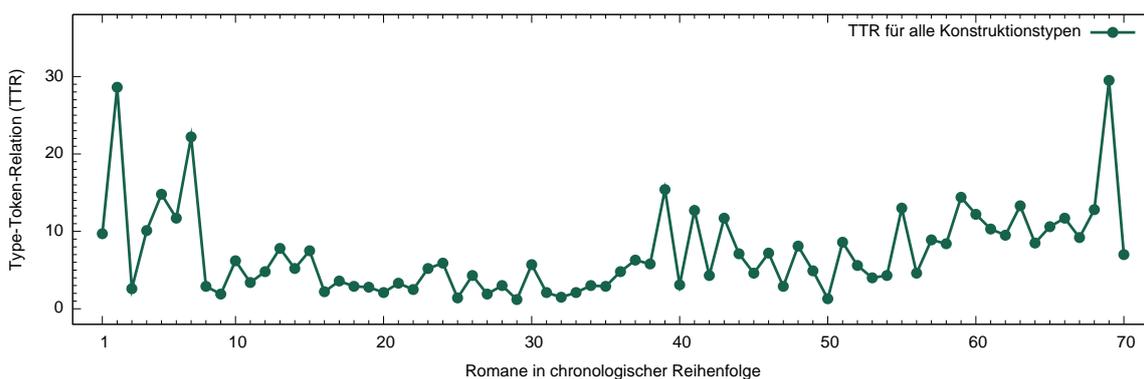


Abbildung 4.29: Type-Token-Relation (TTR) der *verba communicandi* für die gesamte Redebegeleitung, unabhängig vom Konstruktionstyp. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1. Die hohe TTR zu Beginn des Untersuchungszeitraums erklärt sich durch wenige VC-Tokens, die sich auf wenige VC-Types verteilen. Ähnlich verhält es sich auch für das letzte Drittel des Untersuchungszeitraums, wo sich durch Abnahme der direkten Rede einerseits sowie Reduzierung der Redebegeleitung zugunsten reedeinleitender und alleinstehender Turns andererseits wenige VC-Tokens auf eine über die Jahre hinweg relativ konstant gebliebene Anzahl von VC-Types verteilt (siehe auch die Abbildungen 4.1 und 4.28).

Ab 1890, d.h. ab Text Nr. 59²⁰⁸, ist für fast alle in diesem Zeitraum entstandenen Romane eine deutliche Zunahme der TTR zu erkennen. Von den Texten 59-70 wurde nur für vier von zwölf Texten (Nr. 62, 64, 67, 70) eine TTR von weniger als 10% errechnet, während die TTR für die restlichen sieben Romane zweistellig ist und sich auf zwischen 10,3% und 29,5%²⁰⁹ beläuft. Dass die TTR im letzten Drittel des Untersuchungszeitraums allmählich zunimmt, hat mehrere Gründe. Zum einen nimmt die direkte Rede im Vergleich zum Erzähltext ab (erkennbar in Abbildung 4.2), während die direkte Rede gleichzeitig seltener in Verbindung mit einer Redebegeleitung steht. Stattdessen werden Turns vermehrt entweder eingeleitet oder sie stehen unbegleitet (siehe Abbildung 4.4). Ein geringerer Anteil von direkter Rede, verbunden mit einem Rückgang der Redebegeleitung, bedeutet wenige Dialogstellen bzw. wenige VC-Tokens (vgl. Abbildung 4.3). Da die Anzahl der VC-Types jedoch nicht analog dazu abnimmt, sondern weitgehend konstant bleibt (vgl. blaue Kurve in Abbildung 4.28), steigt die TTR. Somit ist die TTR der *verba communicandi* ab 1862 bzw. Text Nr. 39, vor allem aber für das letzte Korpusdrittel, meist höher als noch in den Texten Nr. 8-38. Dieser Zusammenhang wird besonders deutlich

²⁰⁸ Émile Zola, *La bête humaine* (1890).

²⁰⁹ Colette, *Chéri*, (1920). In diesem Zusammenhang sei auch auf Abbildung 4.13 verwiesen, wo neben den Standard-VC-Types viele weitere VC-Types enthalten sind (siehe grauer Bereich). Dies deutet bereits auf eine hohe TTR hin.

anhand der Romane Nr. 55²¹⁰, 59²¹¹, 63²¹², 68²¹³ und, wie bereits erwähnt, Nr. 69²¹⁴. Mit Ausnahme des letztgenannten sind die Anteile der direkten Rede bzw. die Anteile der Redebegleitung und letztlich auch die TTR in allen diesen Texten ähnlich hoch.

Zusätzlich zur Betrachtung der TTR für alle VC-Tokens und VC-Types in der Redebegleitung lohnt sich auch ein Blick auf die TTR der *verba communicandi* für die hochfrequenten Konstruktionstypen T1, T2a, T3a, T3b, T3c und T3e²¹⁵, deren Entwicklung durch die Abbildungen 4.30 und 4.31 graphisch dargestellt wird.

Da der **Konstruktionstyp T1** (« ... », *dit-il.*) in nahezu allen Romanen innerhalb der Redebegleitung der am häufigsten verwendete ist²¹⁶, ähnelt der Kurvenverlauf der TTR für diesen Typ (vgl. Abbildung 4.30(a)) sehr dem der TTR für alle Konstruktionstypen (vgl. Abbildung 4.29). Demzufolge gibt es eine frühe Phase (Texte Nr. 1-7) mit einer erhöhten TTR zwischen 9,8% und 33,3%, die durch vergleichsweise wenig direkte Rede, daraus resultierenden wenigen VC-Tokens und einem begrenzten Vorrat an *verba communicandi* gekennzeichnet ist. Von 1823 bis 1850 (Texte Nr. 8-37) ist die TTR mit meist weniger als 10%²¹⁷, vergleichsweise niedrig, und auch die Schwankungen sind moderat. Für eine kurze Phase, die von 1857-1869 dauert (Texte Nr. 38-44), ändert sich dies, indem die Schwankungen durch Zu- und Abnahme der TTR-Werte größer werden. Ab 1869 bis zum Ende des Untersuchungszeitraums (Texte Nr. 45-70) liegen die TTR-Werte wieder enger beieinander und fallen trotz des Minimums von 1,1%²¹⁸ und des Maximums von 41,5%²¹⁹ insgesamt höher aus als zuvor.

Der Verlauf der TTR für die übrigen Konstruktionstypen sieht hingegen anders aus. Für T2a, T3a, T3b und T3c beträgt die TTR vereinzelt auch 0% und 100%, was vor allem auf die Werke Voltaires und Bernardin de Saint-Pierres (vgl. Abbildungen 4.30(b), 4.30(c), 4.31(a) und 4.31(b)) zutrifft. Eine TTR von 0% liegt dann vor, wenn in dem betreffenden Text überhaupt keine VC-Tokens für den jeweiligen Konstruktionstyp ermittelt

²¹⁰ Gustave Flaubert, *Bouvard et Pécuchet* (1881). Anteil der direkten Rede: 14%, Anteil der Redebegleitung: 21%, TTR: 13%.

²¹¹ Émile Zola, *La bête humaine* (1890). Anteil der direkten Rede: 17,9%, Anteil der Redebegleitung: 24%, TTR: 14,4%.

²¹² Émile Zola, *Le docteur Pascal* (1893). Anteil der direkten Rede: 21,2%, Anteil der Redebegleitung: 25%, TTR: 13,3%.

²¹³ Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918). Anteil der direkten Rede: 24%, Anteil der Redebegleitung: 28%, TTR: 12,8%.

²¹⁴ Colette, *Chéri* (1920).

²¹⁵ Eine Übersicht über die einzelnen Konstruktionstypen ist auf Seite 113 abgebildet.

²¹⁶ Dies verdeutlicht Tabelle C.1 in Anhang B.

²¹⁷ In 28 von 30 Texten beläuft sich die TTR auf weniger als 10%, in immerhin 17 von 30 bleibt sie sogar unter der Fünf-Prozent-Marke

²¹⁸ Nr. 50: Jules Verne, *L'île mystérieuse* (1875).

²¹⁹ Nr. 69: Colette, *Chéri* (1920).

wurden. Ein Wert von 100% wird hingegen erreicht, wenn auf eine bestimmte Anzahl VC-Tokens genauso viele VC-Types kommen. Für den Basistyp T1 sowie die Dativkonstruktion T3e kann dieser Fall nicht eintreten, weil beide Konstruktionstypen in jedem Korpustext so häufig und damit so viele VC-Tokens enthalten sind, dass die Anzahl der VC-Types zu keinem Zeitpunkt auch nur annähernd ähnlich hoch sein kann. Liegt für einen Konstruktionstyp eine TTR von 100% vor, ist dies – zumindest für die im Arbeitskorpus enthaltenen Texte – ein klares Indiz dafür, dass er in dem betreffenden Roman schwach frequent ist. In insgesamt 18 der 70 Romane beträgt die TTR für einen oder mehrere der Konstruktionstypen T2a, T3a, T3b und T3c 100 Prozent. Davon sind allein sechs Texte von Voltaire bzw. Bernardin de Saint-Pierre betroffen, einer von Honoré de Balzac, und die restlichen elf Romane verteilen sich auf den Zeitraum 1862 bis 1922.

Von allen in diesem Kapitel besprochenen Konstruktionstypen weist **Konstruktionstyp T2a** (Konstruktion mit Adjektiv oder *participe passé*) den unregelmäßigsten Kurvenverlauf für die TTR auf. Dass dieser Typ zu den selteneren gehört²²⁰, spiegelt sich auch im TTR-Kurvenverlauf wider. So beläuft sich die TTR der *verba communicandi* (vgl. Abbildung 4.30(b)) für vier Texte²²¹ auf jeweils 0%, d.h. dieser Konstruktionstyp kommt in diesen Texten innerhalb der Redebegleitung gar nicht vor. Obwohl dies für die Werke Voltaires und de Saint-Pierres nicht weiter erstaunt, ist es für Text Nr. 39, *Salammbô* (1862) von Gustave Flaubert, angesichts des viel späteren Erscheinungszeitpunktes umso ungewöhnlicher. In insgesamt neun Romanen²²² beläuft sich die TTR auf 100%. Abgesehen von den Extrema schwankt die TTR in den Texten Nr. 1-15 (1747-1835) zwischen 23,2%²²³ und 75%²²⁴. Für den Zeitraum 1837-1865 (Texte Nr. 16-42) allerdings ist die TTR insgesamt etwas niedriger, d.h. sie bewegt sich mit Ausnahme des Minimal- und Maximalwerts von 0%²²⁵ bzw. 75%²²⁶ zwischen 7,3%²²⁷ und 60%²²⁸. Nach 1867 bis zum Ende des Untersuchungszeitraums (Texte Nr. 43-70) nimmt die TTR grundsätzlich

²²⁰ Die Frequenzwerte für T2a liegen in der Redebegleitung zwischen 0% und 9,4% (vgl. Abbildung 4.16).

²²¹ Nr. 5: Voltaire, *L'ingénu* (1767), Nr. 6: Voltaire, *La princesse de Babylone* (1768), Nr. 7: Bernardin de Saint-Pierre, *Paul et Virginie* (1787), Nr. 39: Gustave Flaubert, *Salammbô* (1862).

²²² Nr. 1: Voltaire, *Zadig* (1747), Nr. 4: Voltaire, *Le blanc et le noir* (1764), Nr. 13: Honoré de Balzac, *Le médecin de campagne* (1833), Nr. 44: Gustave Flaubert, *L'éducation sentimentale* (1869), Nr. 50: Jules Verne, *L'île mystérieuse* (1875), Nr. 55: Gustave Flaubert, *Bouvard et Pécuchet* (1881), Nr. 59: Émile Zola, *La bête humaine* (1890), Nr. 60: Émile Zola, *L'argent* (1891), Nr. 64: Marcel Prévost, *Les demi-vierges* (1894).

²²³ Nr. 9: Stendhal, *Le rouge et le noir* (1830).

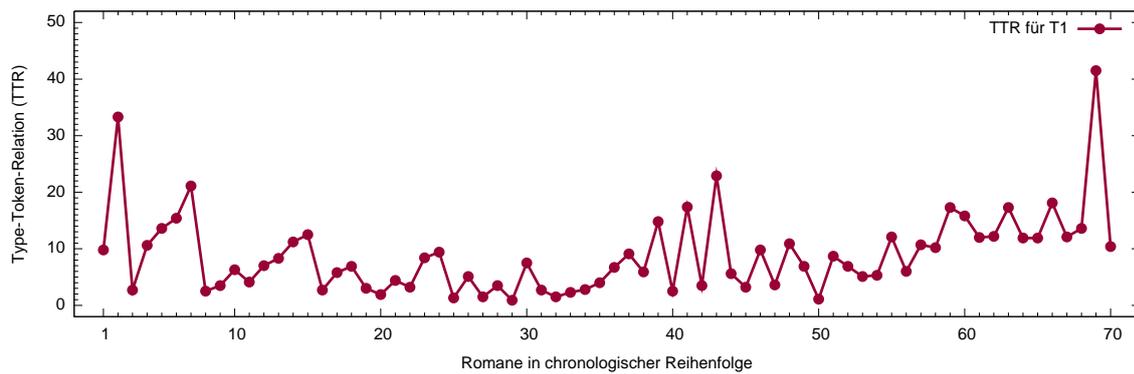
²²⁴ Nr. 15: Honoré de Balzac, *La fille aux yeux d'or* (1835).

²²⁵ Nr. 39: Gustave Flaubert, *Salammbô* (1862).

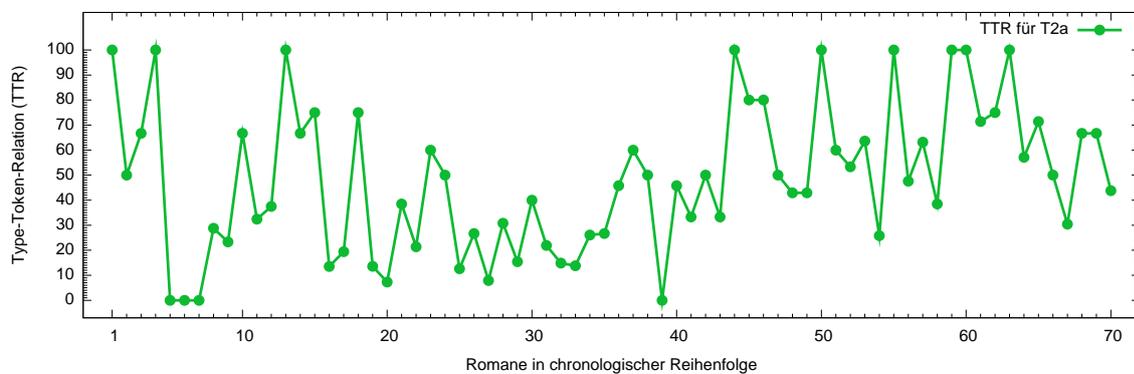
²²⁶ Nr. 18: Honoré de Balzac, *La femme de trente ans* (1842). 4 VC-Tokens verteilen sich auf 3 VC-Types.

²²⁷ Nr. 20: Eugène Sue, *Les mystères de Paris* (1842). Für T2a stehen sich 109 VC-Tokens und 8 VC-Types gegenüber.

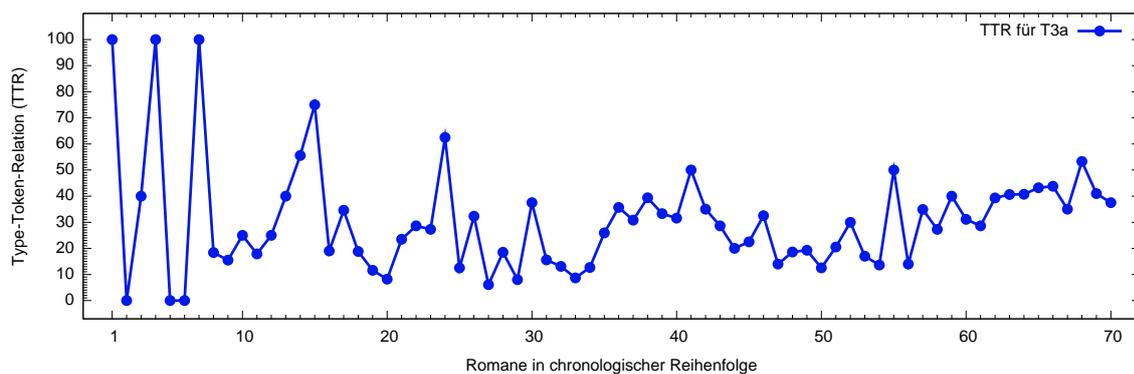
²²⁸ Nr. 23: Honoré de Balzac, *Le colonel Chabert* (1844). T2a enthält 5 VC-Tokens und 3 VC-Types.



(a) TTR der *verba communicandi* für den Konstruktionstyp T1 (« ... », *dit-il.*). Der Kurvenverlauf entspricht im Wesentlichen demjenigen für die gesamte Redebegeleitung (s. Abbildung 4.29).

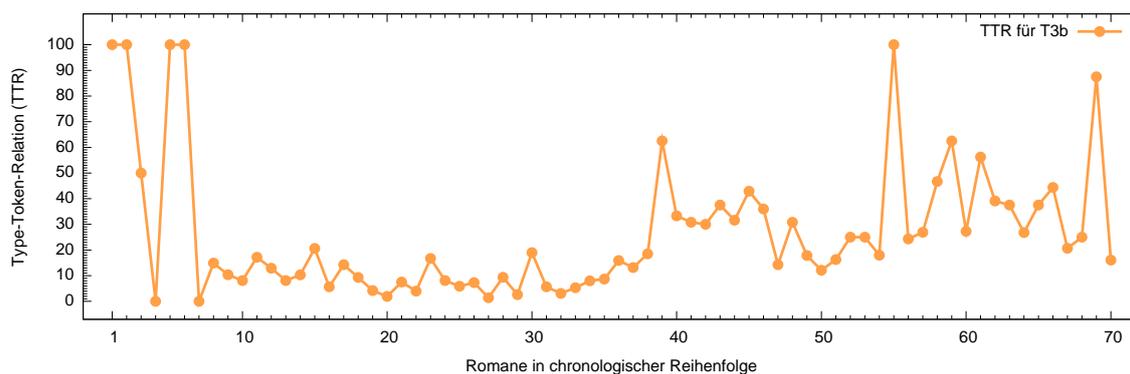


(b) TTR für den Konstruktionstyp T2a (« ... », *répliqua-t-elle stupéfaite.*). Die zwischen den Extremen schwankende Kurve kommt zustande durch wenige VC-Tokens für diesen Typ, die sich meist auf ebenfalls wenige VC-Types verteilen.

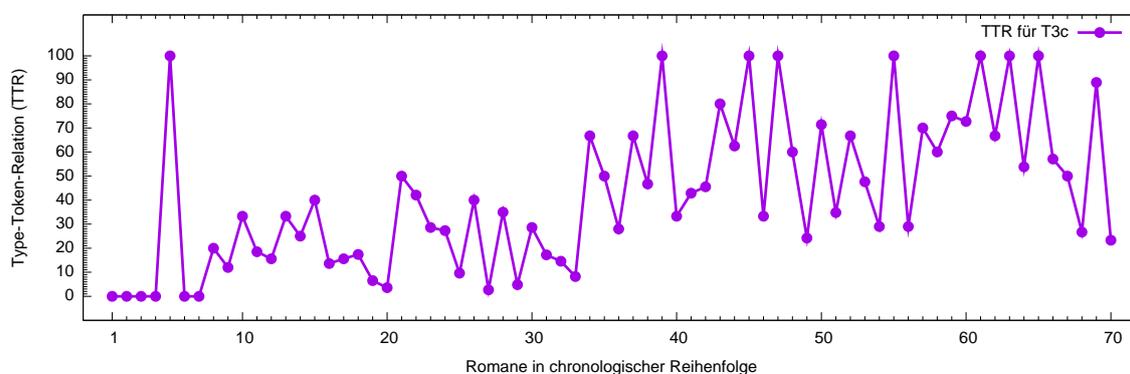


(c) TTR für den Konstruktionstyp 3a (« ... », *s'écria vivement le jeune homme.*).

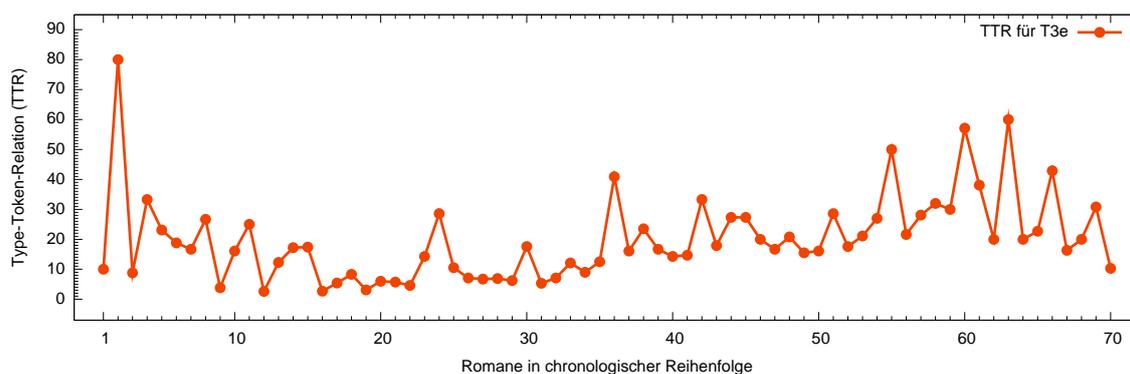
Abbildung 4.30: Type-Token-Relationen (TTR) für die Konstruktionstypen T1, T2a und T3a. Die Aufschlüsselung der Nummern 1-70 für die Romane entlang der horizontalen Achse erfolgt über Tabelle 3.1. Eine TTR von 0% bedeutet, dass gar kein Token für diesen Konstruktionstyp ermittelt wurde. Eine TTR von 100% liegt vor, wenn ebenso viele VC-Types wie VC-Tokens vorhanden sind.



(a) TTR für den Konstruktionstyp T3b (« ... », *remarquait-il en souriant*).



(b) TTR für den Konstruktionstyp T3c (« ... », *répondit-elle avec quelque impatience*). Neben der TTR-Kurve für T2a die am stärksten oszillierende Kurve.



(c) TTR für den Konstruktionstyp T3e (« ... », *dit-elle à son mari*).

Abbildung 4.31: Type-Token-Relationen (TTR) für die Konstruktionstypen T3b, T3c und T3e. Die Aufschlüsselung der Nummern 1-70 für die Romane entlang der horizontalen Achse erfolgt über Tabelle 3.1. Eine TTR von 0% bedeutet, dass gar kein Token für diesen Konstruktionstyp ermittelt wurde. Eine TTR von 100% liegt vor, wenn ebenso viele VC-Typen wie VC-Tokens vorhanden sind.

wieder zu, jedoch werden auch die Ausschläge nach oben und unten wieder größer, so dass die TTR zwischen 25,8%²²⁹ und 100%²³⁰ oszilliert.

Für den **Konstruktionstyp T3a** (Adverbialkonstruktion) weisen außer *Candide* (1759) nur die sehr frühen Texte Nr. 1-7 von Voltaire und Bernardin de Saint-Pierre extreme Werte hinsichtlich der TTR der *verba communicandi* auf (vgl. Abbildung 4.30(c)). So wurde für die Texte Nr. 2, 5 und 6²³¹ keine einzige Adverbialkonstruktion innerhalb der Redebegleitung gezählt, d.h. die TTR beträgt jeweils 0%, während für die Romane Nr. 1, 4 und 7²³² jeweils eine TTR von 100% errechnet wurde. Zwischen 1823 und 1850 (Texte Nr. 8-35) sinkt die TTR-Kurve auf ein deutlich niedrigeres Niveau, d.h. die TTR bewegt sich zwischen 6,1%²³³ und 62,5%²³⁴, jedoch ist der Verlauf durch ein häufiges Auf und Ab gekennzeichnet. Nach 1850 (Texte Nr. 36-70) stabilisiert sich der Kurvenverlauf, so dass es nur wenige extreme Ausreißer nach oben oder unten gibt. Bis 1883 (Nr. 56) nehmen die TTR-Werte außerdem leicht ab. Für den Zeitraum 1888-1922 kann man wieder von einer leichten Zunahme der TTR sprechen, verbunden mit einem vergleichsweise ausgeglichenen Kurvenverlauf mit Werten zwischen 27,3%²³⁵ und 53,3%²³⁶.

Die Annahme, dass eine niedrige TTR einhergeht mit einem großen Anteil von direkter Rede (vgl. Abbildung 4.1), vielen VC-Tokens innerhalb der Redebegleitung und einer mittleren Anzahl von VC-Types (vgl. Abbildung 4.28), bestätigt sich anhand der Texte Nr. 20²³⁷, 27²³⁸, 29²³⁹ und 33²⁴⁰, genauso wie der Umkehrschluss, dass ein hoher Prozentsatz zustande kommt durch vergleichsweise wenige VC-Tokens, gegebenenfalls bedingt durch einen geringen Anteil von direkter Rede (vgl. Abbildung 4.1), aber mittelgroßer

²²⁹ Nr. 54: Émile Zola, *Nana*. Für T2a stehen sich 31 VC-Tokens und 8 VC-Types gegenüber.

²³⁰ Siehe auch Fußnote Nr. 222.

²³¹ Nr. 2: Voltaire, *Memnon* (1750). Nr. 5: Voltaire, *L'ingénu* (1767). Nr. 6: Voltaire, *La princesse de Babylone* (1768).

²³² Nr. 1, Voltaire, *Zadig* (1747): 2 VC-Types für 2 VC-Tokens. Nr. 4, Voltaire, *Le blanc et le noir* (1764): 1 VC-Type für 1 VC-Token. Nr. 7, Bernardin de Saint-Pierre, *Paul et Virginie* (1787): 1 VC-Type für 1 VC-Token.

²³³ Nr. 27: Eugène Sue, *Le juif errant* (1845): 461 VC-Tokens werden aus 28 VC-Types gebildet.

²³⁴ Nr. 24: Honoré de Balzac, *Le lys dans la vallée* (1844). 8 VC-Tokens werden aus 5 VC-Types gebildet.

²³⁵ Nr. 58: Émile Zola, *La terre* (1887). 66 VC-Tokens verteilen sich auf 18 VC-Types.

²³⁶ Marcel Proust, *à l'ombre des jeunes filles en fleurs* (1918).

²³⁷ Eugène Sue, *Les mystères de Paris* (1842): 27 VC-Types für 330 VC-Tokens (T3a) bei insgesamt 3631 VC-Tokens und 75 VC-Types und einer TTR von 2,1% für die gesamte Redebegleitung.

²³⁸ Eugène Sue, *Le juif errant* (1845): 28 VC-Types für 461 VC-Tokens (T3a) bei insgesamt 4501 VC-Tokens und 85 VC-Types und einer TTR von 1,9% für die gesamte Redebegleitung.

²³⁹ Alexandre Dumas, *Le comte de Monte Christo*: 18 VC-Types für 226 VC-Tokens (T3a) bei insgesamt 5821 VC-Tokens und 67 VC-Types und einer TTR für die gesamte Redebegleitung von 1,2%.

²⁴⁰ Alexandre Dumas, *Le collier de la reine* (1849): 22 VC-Types für 254 VC-Tokens (T3a) bei insgesamt 3017 VC-Tokens und 63 VC-Types und einer TTR von 2,1% für die gesamte Redebegleitung.

Zahl von VC-Types, wie es für die Texte 15²⁴¹ und 24²⁴² der Fall ist.

Auch der **Konstruktionstyp T3b** (*gérondif*-Konstruktion) ist in den frühen Texten 1-7 entweder gar nicht enthalten, d.h. die TTR (vgl. Abbildung 4.31(a)) beträgt 0% (Texte Nr. 4 und 7²⁴³), oder aber die TTR beläuft sich auf 100%, d.h. in den Texten Nr. 1, 2, 5 und 6²⁴⁴ sind jeweils genauso viele VC-Tokens wie VC-Types enthalten. Lediglich Text Nr. 3, *Candide* (1759) von Voltaire, fällt diesbezüglich mit einer TTR von 50%²⁴⁵ aus dem Rahmen. Die Redebegleitung für den Zeitraum 1823 bis 1857 (Texte Nr. 8-38) enthält überdurchschnittlich viele *gérondif*-Konstruktionen²⁴⁶, was in diesem Fall auch die TTR-Kurve der *verba communicandi* durch niedrige TTR-Werte ausdrückt. Vor allem bei Honoré de Balzac, George Sand und Eugène Sue sind die Werte niedrig, d.h. sie bewegen sich zwischen 1,5%²⁴⁷ und 21,9%²⁴⁸ und schwanken im Vergleich zu anderen Konstruktionstypen nur moderat. Eine besonders niedrige TTR haben in diesem Fall u.a. vier Texte (Nr. 20²⁴⁹, 27²⁵⁰, 29²⁵¹ und 32²⁵²), von denen die drei erstgenannten durch ein ähnliches Verhalten für die Adverbialkonstruktion (T3a) auffallen²⁵³. Ab dem Erscheinen von *Salammô* von Gustave Flaubert (Nr. 39) im Jahre 1862 ist der Kurvenverlauf der TTR bis zum Ende des Untersuchungszeitraums sehr viel unruhiger. Die TTR der verbleibenden 32 Texte liegt für 26 meist deutlich über 20% und weist einzelne Ausschläge nach oben auf, wie es z.B. die Texte Nr. 39²⁵⁴, 55²⁵⁵ (100%) und 69²⁵⁶ (85,7%) belegen. Für den Zeitraum 1862 bis 1871 (Texte Nr. 40-46) steigt die TTR kurzfristig an, sinkt jedoch zwischen 1873 und 1879 (Texte Nr. 47-54) wieder auf Werte unter 20%. Im letzten

²⁴¹ Honoré de Balzac, *La fille aux yeux d'or* (1835): 3 VC-Types für 4 VC-Tokens (T3a) bei insgesamt 160 VC-Tokens und 12 VC-Types, was einer TTR von 7,5% für die gesamte Redebegleitung entspricht.

²⁴² Honoré de Balzac, *Le lys dans la vallée* (1844): 8 VC-Tokens für 5 VC-Types (T3a) bei insgesamt 288 VC-Tokens und 17 VC-Types, was eine TTR von 5,9% für die gesamte Redebegleitung ergibt.

²⁴³ Voltaire, *Le blanc et le noir* (1764) bzw. Bernardin de Saint-Pierre, *Paul et Virginie* (1787).

²⁴⁴ Voltaire, *Zadig* (1747), Voltaire, *Memnon* (1750), Voltaire, *L'ingénu* (1767) bzw. Voltaire, *La Princesse de Babylone* (1768).

²⁴⁵ 1 VC-Type für 2 VC-Tokens.

²⁴⁶ Vergleiche Abbildung 4.18.

²⁴⁷ Eugène Sue, *Le juif errant* (1845).

²⁴⁸ Honoré de Balzac, *La fille aux yeux d'or* (1835).

²⁴⁹ Eugène Sue, *Les mystères de Paris* (1842): 14 VC-Types für 702 VC-Tokens (T3b) bei insgesamt 3631 VC-Tokens und 75 VC-Types und einer TTR von 2,1% für die gesamte Redebegleitung.

²⁵⁰ Eugène Sue, *Le juif errant* (1845): 12 VC-Types für 811 VC-Tokens (T3b) bei insgesamt 4501 VC-Tokens und 85 VC-Types und einer TTR von 1,9% für die gesamte Redebegleitung.

²⁵¹ Alexandre Dumas, *Le comte de Monte Christo* (1846): 15 VC-Types für 591 VC-Tokens (T3b) bei insgesamt 5821 VC-Tokens und 67 VC-Types und einer TTR für die gesamte Redebegleitung von 1,2%.

²⁵² Honoré de Balzac, *Cousine Bette* (1848): 11 VC-Types für 338 VC-Tokens (T3b) bei insgesamt 1583 VC-Tokens und 23 VC-Types und einer TTR von 2,1% für die gesamte Redebegleitung.

²⁵³ Vgl. die Fußnoten 237, 238 und 239.

²⁵⁴ Gustave Flaubert, *Salammô* (1862).

²⁵⁵ Gustave Flaubert, *Bouvard et Pécuchet* (1881).

²⁵⁶ Colette, *Chéri* (1920).

Viertel des Untersuchungszeitraums ergeben sich die größten Schwankungen von Text zu Text. Die TTR weist vor allem zum Ende hin einen leichten Abwärtstrend auf und liegt für den letzten Text, *Le côté de Guermantes* (1922) von Marcel Proust knapp unter der 20%-Marke.

Der Verlauf der TTR-Kurve für den **Konstruktionstyp T3c** (Präpositionalsyntagma) unterliegt während des gesamten Untersuchungszeitraums, ähnlich wie die TTR-Kurve für den Konstruktionstyp T2a, großen Schwankungen (vgl. Abbildung 4.31(b)). In einer ersten frühen Phase von 1747-1787 (Texte Nr. 1-7) ist der Konstruktionstyp T3c nur in *L'ingénu* (1767) von Voltaire mit einer TTR von 100%, die durch ein einziges VC-Token bzw. einen VC-Type zustande kommt, vorhanden. Eine weitere Phase lässt sich für den Zeitraum 1823 bis 1849 (Texte Nr. 8-33) ausmachen, innerhalb derer die TTR-Werte für T3c mit zwischen 2,7%²⁵⁷ und 50%²⁵⁸ deutlich schwanken. Diese Bewegung setzt sich während einer weiteren, deutlich größeren Phase zwischen 1850 und 1922 fort, wenn auch auf einem deutlich höheren Niveau, d.h. mit TTR-Werten zwischen 23,3%²⁵⁹ und 100%²⁶⁰.

Die TTR-Kurve für den **Konstruktionstyp T3e** (Dativkonstruktion) weist, ähnlich wie Konstruktionstyp T1, wenige extreme Schwankungen auf (vgl. Abbildung 4.31(c)). Für die frühen Texte von Voltaire und Bernardin de Saint-Pierre fällt zunächst auf, dass dieser Konstruktionstyp dort recht häufig in der Redebegleitung anzutreffen ist – erkennbar an TTR-Werten von 33,3%²⁶¹ und niedriger. Einen Ausreißer nach oben stellt die TTR für Text Nr. 2, *Memnon* (1750) von Voltaire, dar, die 80% beträgt und in keinem anderen Text übertroffen wird. Bis 1837 bewegen sich die Werte auf und ab – für die Texte Nr. 9²⁶² und 12²⁶³ liegen sie bei nur 3,8% bzw. 2,6% –, bevor dann von 1837-1850 (Texte Nr. 16-35) eine Phase mit recht konstantem, leicht ansteigendem Kurvenverlauf folgt²⁶⁴. Dieser Trend setzt sich für den Zeitraum 1850-1890 (Texte Nr. 36-59) fort, auch wenn die Werte aufgrund erneuter Schwankungen nicht mehr so eng beieinanderliegen: Die

²⁵⁷ Eugène Sue, *Le juif errant* (1845): 474 VC-Tokens verteilen sich auf 13 VC-Types.

²⁵⁸ Honoré de Balzac, *Eugénie Grandet* (1843): 14 VC-Tokens setzen sich aus 7 VC-Types zusammen.

²⁵⁹ Marcel Proust, *Le côté de Guermantes* (1922).

²⁶⁰ Je einmal in *Salammbô* (1862) von Gustave Flaubert, in *Vingt-mille lieues sous les mers* (1869) und in *Le tour du monde en quatre-vingt jours* (1873) von Jules Verne.

²⁶¹ Voltaire, *Le blanc et le noir* (1764). Für T3c wurden 3 VC-Types für 34 VC-Tokens ermittelt.

²⁶² Stendhal, *Le rouge et le noir* (1830).

²⁶³ George Sand, *Indiana* (1832).

²⁶⁴ Lediglich zwei Texte weichen dabei stärker nach oben hin ab: Text Nr. 24, *Le lys dans la vallée* (1844) von Honoré de Balzac hat eine TTR von 28,6%, und in Nr. 30, *La mare au diable* (1846) von George Sand beläuft sich die TTR auf 17,6%.

TTR für T3e beträgt hier zwischen 14,3% (Nr. 40)²⁶⁵ und 50% (Nr. 55)²⁶⁶. Gegen Ende des Untersuchungszeitraums schwankt die TTR für T3e in den Texten Nr. 60-70 wieder stärker, indem sich die Werte zwischen 10,3%²⁶⁷ und 57,1%²⁶⁸ bewegen und eine leichte Abwärtsbewegung der Kurve anzeigen.

Beim Vergleich der Kurvenverläufe für die Schaubilder 4.30(b), 4.30(c), 4.31(a) und 4.31(b) fällt auf, dass manche Texte gleich für mehrere Konstruktionstypen sehr hohe Werte erreichen. Für *Bouvard et Pécuchet* (1881) von Gustave Flaubert (Text Nr. 55) beträgt die TTR der *verba communicandi* in drei Konstruktionstypen 100%²⁶⁹, was anzeigt, dass der jeweilige Konstruktionstyp zum einen nicht sehr prominent ist und zum anderen ebenso viele VC-Typen wie VC-Tokens vorhanden sind. Ähnlich verhält es sich für *La bête humaine* (1890) von Émile Zola (Text Nr. 59). Die TTR für T2a beläuft sich darin auf 100%, die für T3a liegt bei ca. 40%, für T3c wurde ein Wert von 84% ermittelt, und die TTR für T3b beträgt knapp über 60%²⁷⁰. Eine besondere Position nimmt *Chéri* (1920) von Colette ein. Von allen Romanen hat dieser mit 29,5%²⁷¹ die höchste TTR innerhalb der Redebegleitung (vgl. Abbildung 4.29), gleichzeitig aber korpusweit den zweithöchsten Anteil von direkter Rede²⁷², d.h. 70,8% (vgl. Abbildung 4.1). Dennoch sind von insgesamt 886 nur 227 redebegleitete Dialogstellen (vgl. Abbildung 4.3) mit ebenso vielen VC-Tokens enthalten, von denen die meisten auf die Konstruktionstypen T1 (Basistyp) und T3a (Adverbialkonstruktion) entfallen. Auf alle anderen Konstruktionstypen verteilen sich gerade noch 70 VC-Tokens, was hohe TTR-Werte für die weniger frequenten Konstruktionstypen nach sich zieht, besonders für T2a (Konstruktion mit Adjektiv oder *participe passé*, Abbildung 4.30(b)), T3b (*gérondif-Konstruktion*, Abbildung 4.31(a)) und T3c (Präpositionalsyntax, Abbildung 4.31(b))²⁷³.

Grundsätzlich gilt, dass ein Konstruktionstyp mit einer TTR von 100% in dem be-

²⁶⁵ Victor Hugo, *Les misérables* (1862): 56 VC-Tokens verteilen sich auf 8 VC-Typen.

²⁶⁶ Gustave Flaubert, *Bouvard et Pécuchet* (1881): Auf 2 VC-Typen kommen 4 VC-Tokens.

²⁶⁷ Marcel Proust, *Le côté de Guermantes* (1922): 58 VC-Tokens teilen sich auf 6 VC-Typen auf.

²⁶⁸ Émile Zola, *L'argent* (1891): 7 VC-Tokens werden zu 4 VC-Typen zusammengefasst.

²⁶⁹ Gustave Flaubert, *Bouvard et Pécuchet* (1881): T2a: 4 VC-Tokens/4 VC-Typen; T3b: 1 VC-Token/1 VC-Type; T3c: 1 VC-Token/1 VC-Type.

²⁷⁰ Émile Zola, *La bête humaine* (1890): Konstruktionstyp 2a: 4 VC-Typen für 4 VC-Tokens (100%). Konstruktionstyp 3a: 10 VC-Typen für 25 Tokens (40%). Konstruktionstyp 3b: 3 VC-Typen für 4 VC-Tokens (75%). Konstruktionstyp 3c: 3 VC-Typen für 4 VC-Tokens (75%).

²⁷¹ Dies entspricht 227 redebegleitenden VC-Tokens, die sich auf 67 VC-Typen verteilen. Zum Vergleich: *Memnon* (1750) von Voltaire hat mit 28,6% ebenfalls eine hohe TTR, die sich jedoch aus nur 21 redebegleitenden VC-Tokens und 6 VC-Typen ergibt.

²⁷² Nur in *Le médecin de campagne* (1833) von Honoré de Balzac ist der Anteil von direkter Rede mit 77,5% noch größer. Von 743 Dialogstellen sind immerhin weit mehr als die Hälfte, d.h. 475, in der Redebegleitung anzutreffen, was ebenso vielen VC-Tokens entspricht.

²⁷³ T2a: 8 VC-Typen für 12 VC-Tokens; T3b: 6 VC-Typen für 7 VC-Tokens; T3c: 8 VC-Typen für 9 VC-Tokens.

treffenden Roman zu den seltener verwendeten gehört. Umgekehrt besagt eine niedrige TTR mit weniger als 10%, dass der Konstruktionstyp zu den höherfrequenten zu rechnen ist. Dass es bei diesen Einschätzungen jedoch auch Ausnahmen gibt, zeigt die Type-Token-Relation der *verba communicandi* innerhalb des Konstruktionstyps T3a, d.h. der Adverbialkonstruktion (vgl. 4.30(c)). Besonders gegen Ende des Untersuchungszeitraums (Texte Nr. 57-70) liegt die TTR mit Werten zwischen 27,3% und 53,3% recht hoch, was grundsätzlich den Schluss zuließe, dass T3a im Gegensatz zu den früher entstandenen Korpus-texten, die meist eine deutlich niedrigere TTR aufweisen, zu den weniger frequenten Konstruktionstypen gehören könnte. Ein Blick auf die Abbildung 4.14 zeigt jedoch sofort, dass es sich sogar umgekehrt verhält: T3a ist im letzten Drittel des Untersuchungszeitraums deutlich häufiger vertreten als in davor entstandenen Texten.

Auch bei der in Prozent angegebenen TTR sollte berücksichtigt werden, dass die ihr zugrundeliegenden absoluten Zahlen oftmals aussagekräftiger sind. Eine TTR von 100%, wie sie in insgesamt 18 Romanen für die Konstruktionstypen T2a, T3a, T3b und T3c ermittelt wurde, suggeriert maximale Abwechslung beim Gebrauch von Redeverben, verstellt jedoch auch den Blick auf die Tatsache, dass in keinem dieser Romane mehr als jeweils sechs VC-Tokens bzw. VC-Types²⁷⁴ enthalten sind. Gleiches gilt für TTR-Werte zwischen 50% und 100% herrschen. Aus diesem Grund sollten bei der Beurteilung der TTR immer auch die absoluten Zahlen konsultiert werden.

4.2.5 Refunktionalisierung der *gérondif*-Paraphrase?

Die innerhalb des Arbeitskorpus ab 1823 beobachtbare Zunahme der VC-Types legt die Vermutung nahe, dass sich innerhalb des *discours citant* ein Wandel bezüglich der Redeverben vollzieht. Vivienne Mylne stellte diesbezüglich bereits 1994 eine ab dem 19. Jahrhundert einsetzende Veränderung fest:

.... [E]n dehors des énoncés et des indices d'attribution, on rencontre assez fréquemment des adverbes ou des tournures adverbiales comme *en riant*, *en soupirant*, etc., qui qualifient un *dit-elle* ou un *reprit-il*. Quoi de plus facile que de substituer «soupira-t-elle» à «dit-elle en soupirant»? [...] Ensuite, les romanciers ont profité de leur liberté pour étendre l'usage à beaucoup d'autres verbes, que ceux-ci se réfèrent au son de voix du locuteur – «brailla-t-il» – ou son comportement physique – «minauda-t-elle».²⁷⁵

Ein erstes Indiz für diese Entwicklung beobachtet Vivienne Mylne bei Charles Sorel, der bereits im 17. Jahrhundert *interrompre* als *verbum communicandi* innerhalb der Rede-

²⁷⁴ Gustave Flaubert, *L'éducation sentimentale* (1869).

²⁷⁵ Vivienne Mylne (1994), S. 42.

begleitung verwendet.²⁷⁶ Bis zum 19. Jahrhundert scheint der Großteil der Romanciers keine Notwendigkeit für die Verwendung expressiver *verba communicandi* zu sehen. Jedenfalls kommen in den Erzählungen von Voltaire oder Bernardin de Saint-Pierre nur maximal elf verschiedene *verba communicandi* in der Redebegleitung vor, von denen lediglich *s'écrier* eine Modifikation der Stimme ausdrückt. Erst mit Victor Hugo wird der Einsatz expressiver Redeformen allmählich zum gängigen Stilmittel:

„Pour introduire ces nouveaux verbes déclaratifs, un romancier devait outrepasser l'usage traditionnel et exploiter de nouvelles ressources stylistiques. C'est ce qu'a fait Hugo dans ses romans, tandis que Balzac, moins novateur à cet égard, observait [...] l'usage classique.“²⁷⁷

Für *Notre-Dame de Paris* von Victor Hugo ermittelt Vivienne Mylne (1987) ein Dutzend Verben, die sich auf diese Weise zum Voll-Redeverb entwickelt haben sollen: *balbutier*, *grommeler*, *murmurer*, *bégayer*, *chuchoter*, *glapir*, *hurler*, *jurer*, *nasiller*, *souffler*, *soupirer* und *vociférer*²⁷⁸. Auf das Untermauern dieser These durch weitere Beispiele aus anderen Romanen verzichtet Vivienne Mylne jedoch mit der Begründung, dass sich, obwohl der Untersuchungsgegenstand grundsätzlich von Interesse sei, der immense Aufwand dafür nicht lohne.²⁷⁹ Zu Zeiten, als Untersuchungen dieser Art ausschließlich manuell erfolgten, mag dieser Einwand seine Berechtigung gehabt haben. Das Programm DC-Analyse hält jedoch neben den *verba communicandi* auch die im *discours citant* auftretenden *gérondif*-Konstruktionen fest, so dass die gewonnenen Daten nur noch zum automatisierten Auswerten und Vergleichen aufbereitet werden müssen.

Die Veränderungen sind im Arbeitskorpus klar zu erkennen. Ob sich Vivienne Mylnes Hypothese, die sie auf der Grundlage einzelner Handauszählungen aufgestellt hat, als richtig erweist, lässt sich nun anhand der gewonnenen Daten erstmals quantitativ überprüfen. Soviel vorab: Wenn sich auch die Annahme einer Veränderung bestätigt, so wird die Hypothese, dass vor allem Verben, die innerhalb der Redebegleitung das Redeverb mittels *gérondif*-Konstruktion näher bestimmen, selbst zum Vollredewerb werden (zur Veranschaulichung siehe die folgenden beiden Beispiele), widerlegt.

(11676) – *Oui... je venais... je viens de voir quelqu'un, dit Fleur-de-Marie en balbutiant et en rougissant de honte.*²⁸⁰

(8443) – *Monsieur, balbutia le jeune homme étourdi, [...].*²⁸¹

²⁷⁶ Zitiert nach Vivienne Mylne (1994), S. 42: „Il interrompt X en disant: «Mais non...» devient «Mais non...», interrompt-il.“

²⁷⁷ Vivienne Mylne (1994), S. 42.

²⁷⁸ Vivienne Mylne (1987), S. 415. Sie betont dort ebenfalls: „Hugo is among those novelists who initiated the practice of introducing direct speech with verbs which indicate, by a kind of ellipsis, the manner of speaking: <soupira-t-elle>, for instance, instead of <dit-elle en soupirant>.“ (S. 415).

²⁷⁹ Vgl. Vivienne Mylne (1994), S. 42.

²⁸⁰ Eugène Sue, *Les mystères de Paris* (1842).

²⁸¹ Alexandre Dumas, *Le comte de Monte Christo* (1846).

Die Verbindung aus üblicherweise *dire* als Universal-Redeverb plus *gérondif* wird durch eine „elliptische Reduktion“²⁸² aufgebrochen. Das Verb aus der *gérondif*-Konstruktion tritt an die Stelle von *dire* und schließt sich so unmittelbar an die direkte Rede an. Durch den Wegfall der adverbialen Ergänzung öffnet sich eine Leerstelle, die sich der Schriftsteller für die weitere Ausgestaltung der Figurenrede zunutze machen kann. Dies kann in Gestalt weiterer Ergänzungen des Redeverbs, wie z.B. *étourdi* in Beispiel (8443), geschehen, wo es sich um einen Wechsel von Konstruktionstyp 3b²⁸³ (*gérondif*-Konstruktion) zu Konstruktionstyp 2a (Konstruktion mit Adjektiv oder *participe passé*) handelt. Auf diese Weise ist eine Verstärkung des semantischen Gehalts der Redebegleitung möglich, die den Äußerungen einer Figur noch mehr Nachdruck verleiht und gleichzeitig deren Profil schärft. Jedoch können nicht alle Verben, die Teil der *gérondif*-Paraphrase sind, zu *verba communicandi* werden. Nur wenn zwischen dem Verb aus der Paraphrase und z.B. *dire* als übergeordnetem Redeverb ein „semantisches Inklusionsverhältnis“²⁸⁴ gegeben ist, ist es als auch Redeverb tauglich:

„Das übergeordnete *dire* ist neutral in Bezug auf intonatorische Kennzeichen wie z.B. ‚Tonstärke‘ und ‚Tonqualität‘; solche Merkmale werden durch die Hyponyme spezifiziert.“²⁸⁵

Dies erklärt auch, weshalb z.B. *parler* nicht als Teil einer *gérondif*-Paraphrase eingesetzt werden kann. *Dire* und *parler* befinden sich semantisch auf derselben Ebene, weshalb kein Inklusionsverhältnis vorliegt.

Um zu überprüfen, ob es tatsächlich einen Übergang von der *gérondif*-Konstruktion zum Redeverb gibt, trägt man die jeweilige Anzahl der *gérondif*-Formen und der konjugierten Verbformen gegen die Zeit auf und vergleicht sie miteinander. Genauer gesagt wird zunächst für jedes in der Liste auf S. 45 f. enthaltene *verbum communicandi* ein Schaubild erzeugt, das jeweils die Anzahl der Verb-Tokens sowie die Anzahl der ermittelten *gérondif*-Formen (beide ablesbar an der senkrechten Achse) innerhalb der Redebegleitung für jeden im Korpus enthaltenen Text – diese werden entlang der horizontalen Achse chronologisch angeordnet und durchnummeriert – abbildet. Verbindet man jeweils die Punkte miteinander, ergeben sich zwei Kurven, die die zeitliche Entwicklung illustrieren: eine für die *gérondif*-Formen (blau) und eine für konjugierte Verbformen (rot). Die graphische Darstellung dieser Entwicklung leistet bei der Auswertung wertvolle Hilfe, da Besonderheiten sofort sichtbar werden. Die Datengrundlage für die Abbildungen liefert das Programm DCAnalyse, das bei jedem Programmdurchlauf Listen mit den verwen-

²⁸² Annette Sabban (1988), S. 48.

²⁸³ Eine Übersicht über die Konstruktionstypen ist auf S. 113 f. abgebildet.

²⁸⁴ Vgl. Annette Sabban (1988), S 52.

²⁸⁵ Annette Sabban (1988), S 52.

deten VC-Tokens und *gérondif*-Konstruktionen erzeugt.²⁸⁶

Erläutert wird im Folgenden die Entwicklung derjenigen Verben, deren Kurvenverlauf auffällig ist. Die Anzahl entweder des konjugierten Redeverbs oder der *gérondif*-Form muss in mindestens einem Text sechs oder mehr betragen.²⁸⁷ Das hat jedoch zur Folge, dass die meisten Verben aufgrund ihrer schwachen Frequenz, d.h. weniger als sechs Tokens und der daraus resultierenden geringen Aussagekraft für die Analyse nicht berücksichtigt werden. Von den *verba communicandi*, die die Stimme der sprechenden Figur modifizieren, sind die folgenden elf Verben von Interesse: *balbutier*, *bégayer*, *crier*, *s'écrier*, *grommeler*, *hurler*, *murmurer*, *interrompre*, *soupirer*, *éclater* und *soupirer*.

Die Verwendung von **balbutier** (Abb. 4.32) als Redeverb ist bis 1850 (Nr. 35) eher lückenhaft. Es tritt innerhalb des Korpus erstmalig in *Han d'Islande* (1823) von Victor Hugo (Nr. 8) auf, und zwar viermal ausschließlich als konjugiertes *verbum communicandi*. In einem weiteren seiner Romane, *Notre-Dame de Paris* (1831) (Nr. 11), steht fünf konjugierten Tokens eine *gérondif*-Konstruktion mit *balbutier* gegenüber:

(2546) «Quelle bohémienne à la chèvre ? dit-il en balbutiant.»²⁸⁸

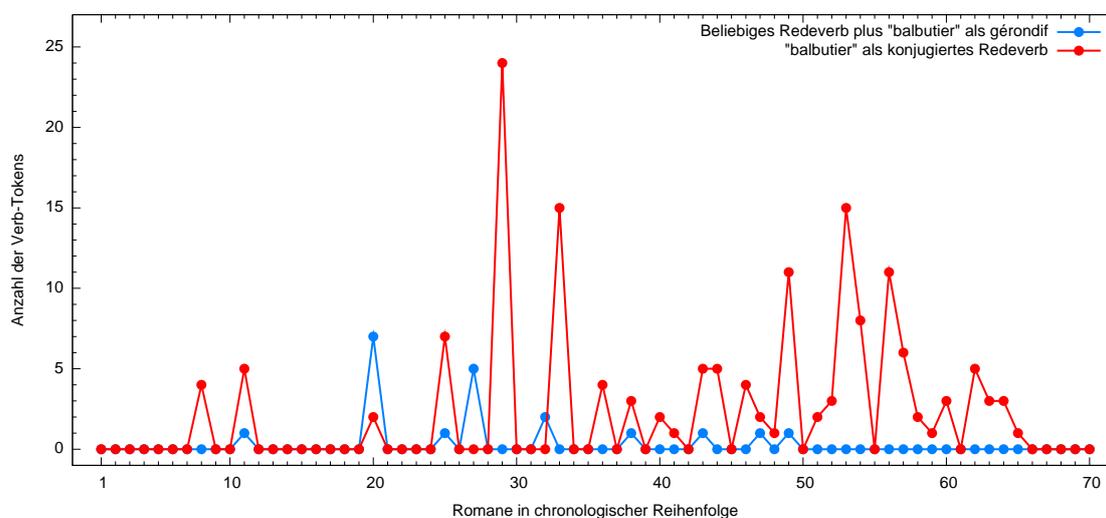


Abbildung 4.32: Entwicklung der konjugierten sowie der *gérondif*-Formen von *balbutier* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

²⁸⁶ Die auszuwertenden Dateien, zwei für jeden Text, sind im Verzeichnis `output/tabellen/` gespeichert und heißen `vd_gerondif_rb_nur_gerondif__dateiname.csv` und `vd_types_rb__dateiname.csv`.

²⁸⁷ Dieser Wert wurde als Mindestgrenze willkürlich festgelegt, da bei der Durchsicht der Schaubilder die Mehrzahl der Verben nicht nur äußerst selten gebraucht wurde, sondern auch bezüglich der besprochenen Problematik überhaupt keine Auffälligkeiten aufwies.

²⁸⁸ Victor Hugo, *Notre-Dame de Paris* (1831).

Erst 1842 ist *balbutier* ein weiteres Mal nachgewiesen, und zwar in *Les mystères de Paris* (1842) (Nr. 20) von Eugène Sue, wo sich das Verhältnis mit sieben *gérondif*-Formen zu nur zwei konjugierten Formen umkehrt, genauso wie in *Le juif errant* (1845) (Nr. 27) mit einem Verhältnis von sechs *gérondif*-Formen zu null konjugierten. Zwischenzeitlich kommt das Verb nur noch bei Alexandre Dumas' *Les trois mousquetaires* (1844) (Nr. 25) vor, allerdings überwiegt hier die konjugierte Form, insgesamt sieben, gegenüber einer *gérondif*-Form. Das Maximum erreicht *balbutier* in Dumas' *Le comte de Monte Christo* (1846) (Nr. 29) und *Le collier de la reine* (1849) (Nr. 33) mit 24 bzw. 15 konjugierten Tokens sowie jeweils null *gérondif*-Formen innerhalb des redebegleitenden *discours citant*. Das folgende Beispiel illustriert, wie die zuvor häufig gebrauchte Kombination *dit-il en balbutiant* aufgebrochen wird. *Balbutier* tritt nicht nur an die Stelle des Hauptverbs, sondern wird außerdem durch Hinzufügen einer *gérondif*-Konstruktion, die auf eine Körperbewegung Bezug nimmt, semantisch erweitert:

(170) – *Cependant, balbutia Dantès en essuyant la sueur qui coulait de son front, cependant je vous avais laissé deux cents francs, il y a trois mois, en partant.*²⁸⁹

Für den Zeitraum zwischen 1850 und 1874 (Nr. 36-49) fällt auf, dass die *gérondif*-Konstruktion nur noch vereinzelt gebraucht wird²⁹⁰. Danach trifft man in der Redebegleitung nur noch auf konjugierte Verbtokens, die vor allem bei Émile Zola und Marcel Prévost vorkommen. Nach 1894 ist *balbutier* weder als Teil einer *gérondif*-Konstruktion noch als konjugiertes Redeverb nachgewiesen.

Für *bégayer* (Abb. 4.33) sieht die Entwicklung anders aus. Als Teil einer *gérondif*-Konstruktion wird es in allen 70 Texten innerhalb der Redebegleitung nur in den ersten 35 zweimal mit je einem Vorkommen eingesetzt: einmal in *Le médecin de campagne* (1833) (Nr. 13) und einmal in *Eugénie Grandet* (1843) (Nr. 21), beide von Honoré de Balzac:

(903) – *Je suis mort, dit-il en bégayant et en faisant un effort affreux pour se dresser.*²⁹¹

(700) – *Nous verrions donc à tâ... tâ... tâcher de nous aaaarranger tou.... tous deux dans les po... po... po... possibilités relatives et sans m'en... m'en... m'engager à quelque chose que je... je... je ne vooooou... oudrais pas faire,*

²⁸⁹ Alexandre Dumas, *Les trois mousquetaires* (1846). Im Roman sind drei weitere Konstruktionen dieser Art enthalten.

²⁹⁰ Je zwei *gérondif*-Tokens wurden bei Gustave Flaubert, *Madame Bovary* (1857) (Nr. 38), Émile Zola, *Thérèse Raquin* (1867) (Nr. 43), Jules Verne, *Le tour du monde* (1873) (Nr. 47) sowie Émile Zola, *La conquête de Plassans* (1874) (Nr. 49) ermittelt.

²⁹¹ Honoré de Balzac, *Le médecin de campagne* (1833).

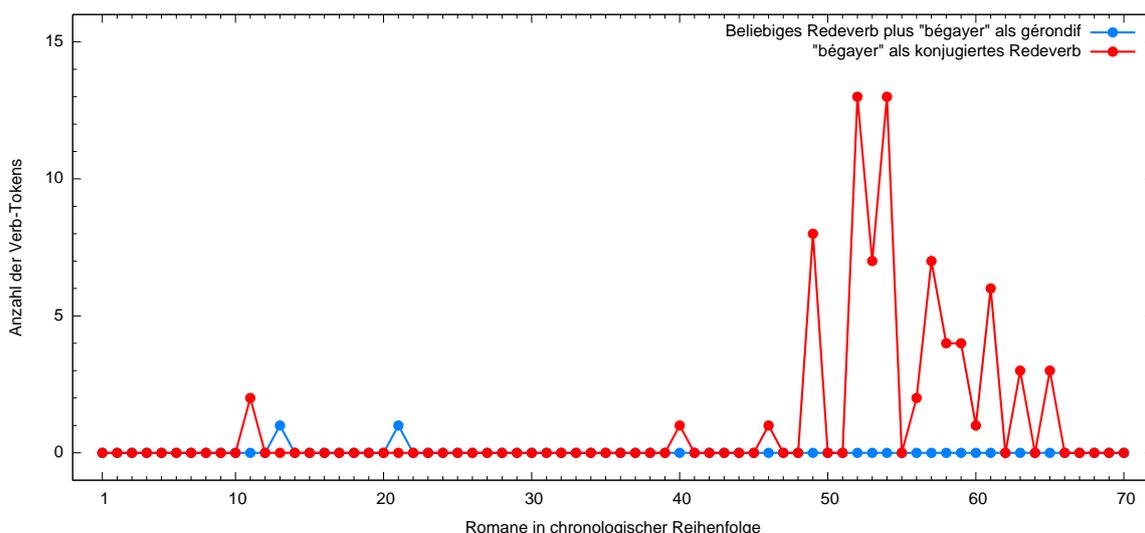


Abbildung 4.33: Entwicklung der konjugierten sowie der *gérondif*-Formen von *bégayer* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

*dit Grandet en bégayant. Parce que, voyez-vous, monsieur le président me demandait naturellement les frais du voyage.*²⁹²

In letzterem Beispiel heben sowohl die *gérondif*-Ergänzung mit *bégayer* als auch die Schreibweise und die Zeichensetzung, d.h. Auslassungspunkte, den stotternden Charakter der Äußerung hervor. Allerdings fällt beim genaueren Lesen auf, dass im zweiten Teil des Turns, der sich an die Redebegleitung anschließt, das Stottern nicht mehr durch eine ähnliche Graphie gekennzeichnet wird. Entweder setzt der Autor stillschweigend voraus, dass der Leser die sprechende Figur weiterhin stotternd wahrnimmt, oder der Sprachfehler liegt tatsächlich nicht mehr vor. Von einem häufigeren Auftreten der konjugierten Form kann man erst ab 1874 sprechen. Bis dahin kommen in drei von 17 Romanen jeweils ein bis zwei konjugierte *bégayer*-Tokens vor, und zwar in *Notre-Dame de Paris* (1831) (Nr. 11) und *Les misérables* (1862) (Nr. 40), beide von Victor Hugo, sowie in *La curée* (1871) (Nr. 46) von Émile Zola, z.B.

(3599) «Oh! *bégaya-t-elle pétrifiée, je savais bien que c'était encore lui!*»²⁹³

(8267) – *Oui! bégaya Marius.*²⁹⁴

(1693) – *Je veux te parler, bégaya-t-elle d'une voix sourde.*²⁹⁵

²⁹² Honoré de Balzac, *Eugénie Grandet* (1843).

²⁹³ Victor Hugo, *Notre-Dame de Paris* (1831).

²⁹⁴ Victor Hugo, *Les misérables* (1862).

²⁹⁵ Émile Zola, *La curée* (1871).

In Beispiel (8267) ist allerdings kaum vorstellbar, dass die sprechende Figur bei der Artikulation eines knappen «*Oui!*» stottert, zumal der Äußerung ein Ausrufezeichen folgt, das ihr einen gewissen Nachdruck verleiht. Nach 1874 (Nr. 49) bleibt die Verwendung von *bégayer* auf die Romane Zolas beschränkt. Zwischen 1871 und 1894 benutzt dieser das Verb als Voll-Redeverb in 13 der 16 von ihm im Korpus enthaltenen Romane. In acht Texten wurden je zwischen vier und 13 konjugierte Verbtokens gezählt.

Die Verben *crier* (Abb. 4.34) und *s'écrier* (Abb. 4.35) zeichnen sich beide dadurch aus, dass sie als Teil einer *gérondif*-Konstruktion nur sehr selten vorkommen, während die konjugierte Verbform zu den am häufigsten verwendeten Redeverben überhaupt gehört.

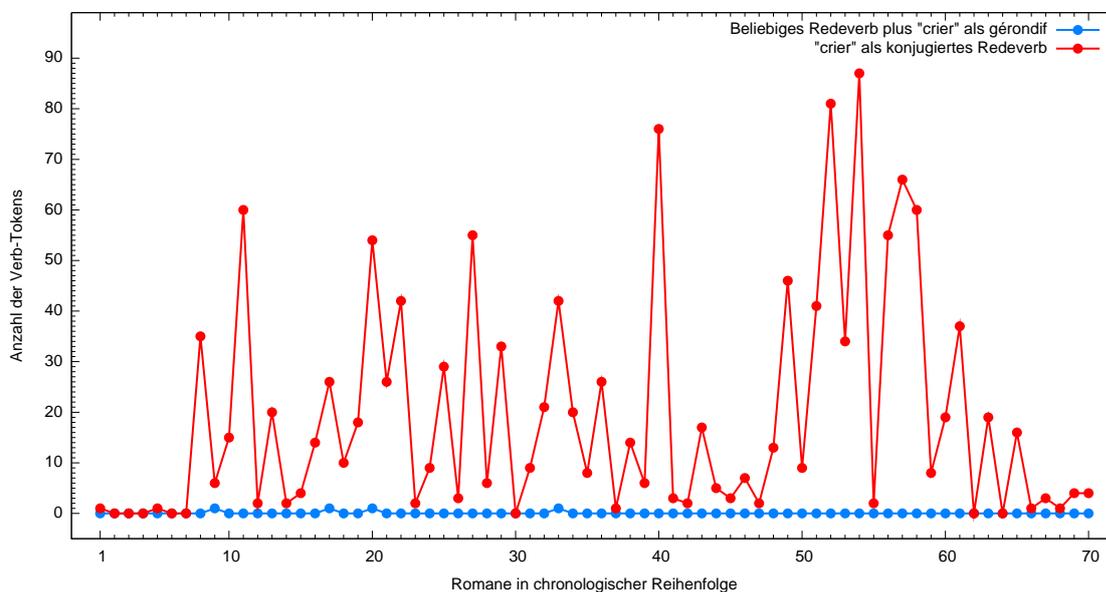


Abbildung 4.34: Entwicklung der konjugierten sowie der *gérondif*-Formen von *crier* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Crier wurde als Teil einer *gérondif*-Konstruktion nur je einmal in vier Romanen ermittelt, und zwar bei Stendhal (*Le rouge et le noir* (1830) (Nr. 9) und *La chartreuse de Parme* (1839) (Nr. 17)), Eugène Sue (*Les mystères de Paris* (1842) (Nr. 20)) und Alexandre Dumas (*Le collier de la reine* (1849) (Nr. 33)), z.B.

(1083) – *Trois cent vingt francs, dit l'autre en criant.*²⁹⁶

Allerdings gibt in dem vorangegangenen Beispiel nur die *gérondif*-Ergänzung den entscheidenden Hinweis auf die Lautstärke des Gesagten. Auf ein zusätzliches Signal wie

²⁹⁶ Stendhal, *Le rouge et le noir* (1830).

ein Ausrufezeichen verzichtet der Autor – was übrigens eine gängige Praxis ist²⁹⁷. Als konjugiertes Verbtoken innerhalb der Redebegeleitung tritt *crier* mit 35 Tokens erstmalig in Victor Hugos *Han d'Islande* (1823) (Nr. 8) häufiger in Erscheinung, z.B.

(342) – *Spiagudry ! cria une voix qui sortait de la salle des cadavres.*²⁹⁸

Zuvor gibt es lediglich je ein Vorkommen in *Zadig* (1747) (Nr. 1) bzw. *L'ingénu* (1767) (Nr. 5) von Voltaire. Die Verteilung von *crier* bis zum Ende des Untersuchungszeitraums ist gekennzeichnet durch ein stetiges Auf und Ab der Tokenanzahl. Nur in *La mare au diable* (1846) (Nr. 30) von George Sand und den beiden Romanen *L'automne d'une femme* (1893) (Nr. 62) und *Les demi-vierges* (1894) (Nr. 64) von Marcel Prévost ist *crier* überhaupt nicht als Redeverb enthalten. Die Spitzenwerte mit über 70 Tokens wurden bei Victor Hugo und Émile Zola ermittelt²⁹⁹.

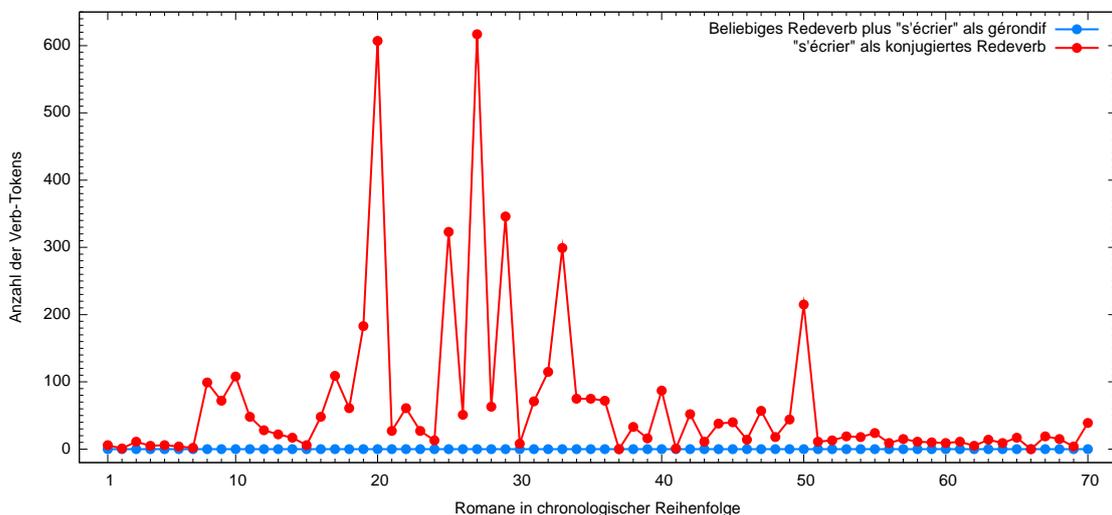


Abbildung 4.35: Entwicklung der konjugierten sowie der *gérondif*-Formen von *s'écrire* innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

S'écrire wird mit Ausnahme dreier Werke³⁰⁰ (Nr. 2, Nr. 37 und Nr. 41) ausschließlich als konjugiertes Redeverb gebraucht und fällt insgesamt durch seine Häufigkeit auf. Die Romane von Eugène Sue sind diesbezüglich die absoluten Spitzenreiter: Wohl auch bedingt durch die enorme Textlänge – die Gesamttokenzahl beläuft sich auf mehr als eine halbe Million Tokens – und dem damit verbundenen Anteil von direkter Rede entfallen

²⁹⁷ Siehe auch Beispiele auf S. 134.

²⁹⁸ Victor Hugo, *Han d'Islande* (1823).

²⁹⁹ Victor Hugo, *Les misérables* (1862) (Nr. 40): 66 Tokens, Emile Zola, *L'assommoir* (1877) (Nr. 52): 81 Tokens und *Nana* (1879) (Nr. 54): 88 Tokens.

³⁰⁰ Voltaire, *Memnon* (1750), George Sand, *François le champi* (1850) und Jules et Édmond de Goncourt, *Germinie Lacerteux* (1865).

auf *Les mystères de Paris* (1831) (Nr. 20) und *Le juif errant* (1845) (Nr. 27) jeweils mehr als 600 *s'écrier*-Tokens. Knapp die Hälfte der im Korpus enthaltenen Texte weist mehr als 50 Tokens auf. Häufungen mit mehr als 200 *s'écrier*-Tokens innerhalb des redebegleitenden *discours citant* konzentrieren sich vor allem auf den Zeitraum 1842 bis 1875. Im letzten Drittel des Untersuchungszeitraums, der viele Romane von Émile Zola enthält, pendelt sich der Kurvenverlauf auf einem deutlich niedrigeren Niveau ein. Hier zeigt sich erneut (vgl. Abbildung 4.12), dass *s'écrier* bei Zola deutlich seltener vorkommt als *crier*.

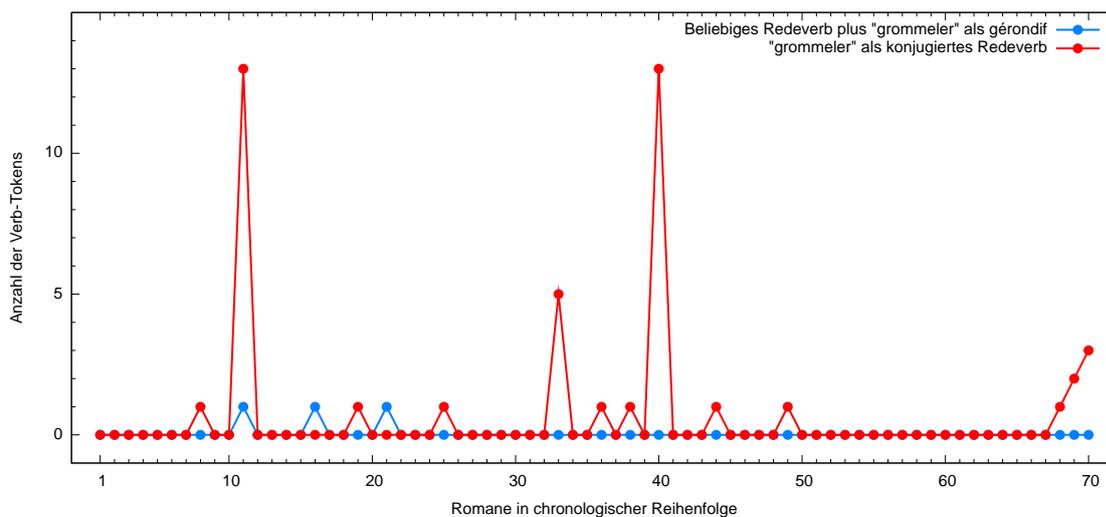


Abbildung 4.36: Entwicklung der konjugierten sowie der *gérondif*-Formen von *grommeler* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Grommeler (Abb. 4.36) ist als *verbum communicandi* in 15 Texten innerhalb des Arbeitskorpus anzutreffen, und zwar fast ausschließlich als konjugierte Verbform. Bei nur je einem Vorkommen in drei Romanen spielen die *gérondif*-Konstruktionen quasi keine Rolle³⁰¹. In 13 Romanen wurden konjugierte Tokens von *grommeler* gefunden, doch kommen nur in dreien davon fünf oder mehr vor. Besonders Victor Hugo verwendet es gerne: Während in *Han d'Islande* (1823) (Nr. 8) nur ein einziges Token enthalten ist, so sind es für *Notre-Dame de Paris* (1831) (Nr. 11) und *Les misérables* (1862) (Nr. 40) je 13, wie z.B.

(1605) – *Deux napoléons ! grommela une vieille édentée qui était là. Qu'en voilà une qui est heureuse !*³⁰²

³⁰¹ Victor Hugo, *Notre-Dame de Paris* (1831) (Nr. 11), Honoré de Balzac, *César Birotteau* (1837) (Nr. 16) und *Eugénie Grandet* (1843) (Nr. 21).

³⁰² Victor Hugo, *Les misérables* (1862).

Alexandre Dumas setzt das Verb in *Le collier de la reine* (1849) (Nr. 33) immerhin fünfmal ein. Zwischen 1875 und 1913, d.h. in den Texten 50-67, tritt *grommeler* gar nicht in Erscheinung. Erst bei Marcel Proust³⁰³ und Colette³⁰⁴ wird es wieder als Redeverb eingesetzt, wenn auch nur in geringer Anzahl.

Das Redeverb ***hurler*** (Abb. 4.37) kommt in insgesamt 19 Texten des Korpus vor, zählt aber wie *grommeler* zu den weniger frequenten Verben. Abgesehen von einem einzigen *gérondif*-Token in *Le comte de monte Christo* (1846) (Nr. 29) von Alexandre Dumas, wird es sonst nur als Voll-Redeverb eingesetzt und ist erstmals bei Victor Hugo in *Notre-Dame de Paris* (1831) (Nr. 11) mit drei konjugierten Verbtokens nachgewiesen:

(14972) – *Mais moi, moi ! s'écria Morrel en hurlant de douleur, moi, je l'aime !*³⁰⁵

(1504) – *Hé bien ! fils, hurlait un perclus en faisant effort pour l'atteindre de sa béquille, nous jetteras-tu encore des sorts du haut des tours de Notre-Dame ?*³⁰⁶

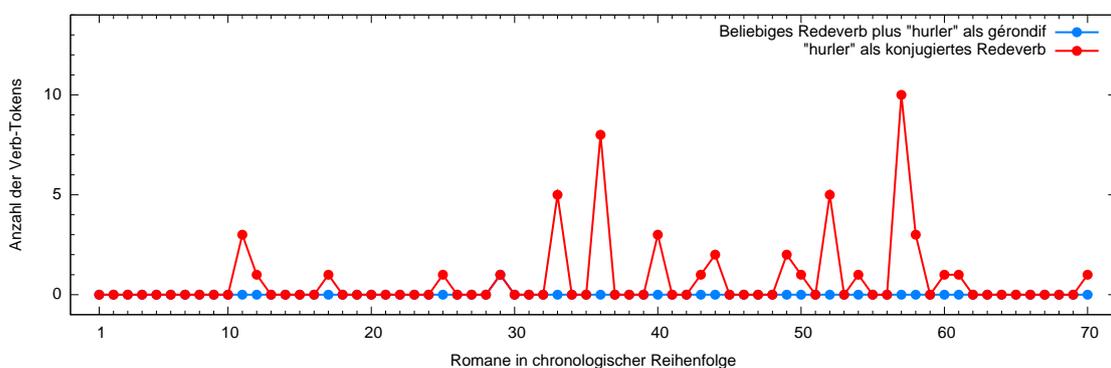


Abbildung 4.37: Entwicklung der konjugierten sowie der *gérondif*-Formen von *hurler* innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Die Frequenz der Tokens variiert jedoch sehr stark: In zehn Romanen findet sich lediglich je ein konjugiertes Token, und nur in vier Romanen sind fünf und mehr Tokens enthalten³⁰⁷. Nach 1890 wird *hurler* jedoch kaum noch eingesetzt und verschwindet allmählich aus dem Redeverb-Repertoire der Schriftsteller.

Mit Werten bis zu 140 konjugierten Verbtokens pro Roman gehört das Redeverb ***murmurer*** (Abb. 4.38) zu den höherfrequenten Redeverben. Auch wenn die *gérondif*-

³⁰³ Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918)(1 Token) und *Le côté de Guermantes* (1922) (3 Tokens).

³⁰⁴ Colette: *Chéri* (1920)(2 Tokens).

³⁰⁵ Alexandre Dumas, *Le comte de Monte Christo* (1846).

³⁰⁶ Victor Hugo, *Notre-Dame de Paris* (1831).

³⁰⁷ Alexandre Dumas, *Le collier de la reine* (1849)(5 Tokens) und *La tulipe noire* (1850)(8 Tokens) sowie Émile Zola, *L'assommoir* (1877)(5 Tokens) und *Germinal* (1885)(10 Tokens).

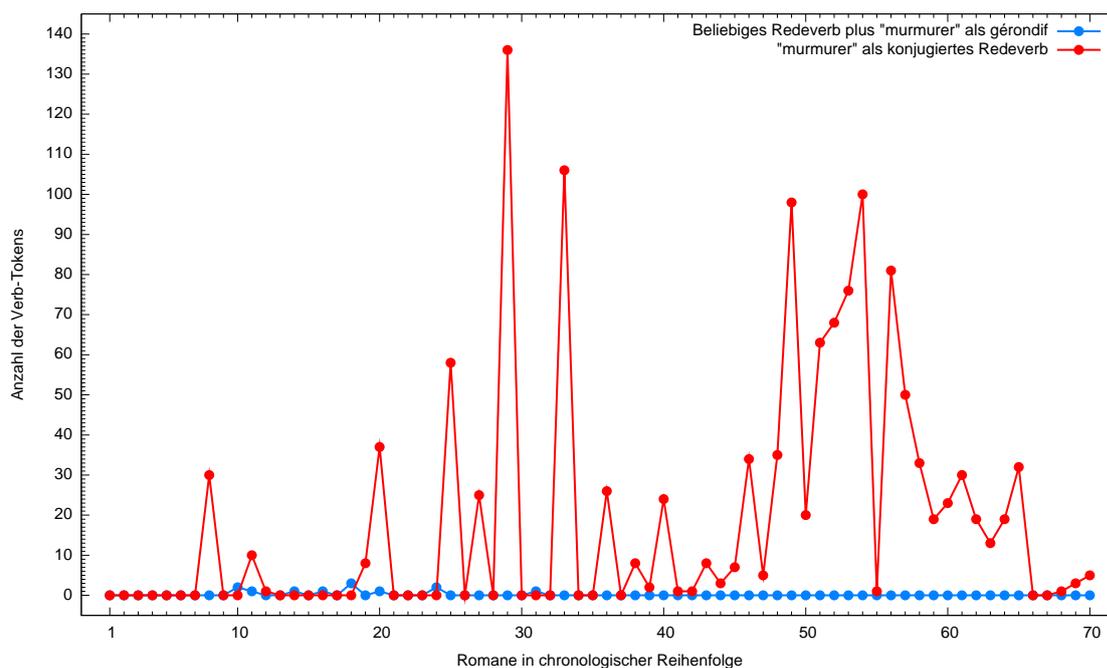


Abbildung 4.38: Entwicklung der konjugierten sowie der *gérondif*-Formen von *murmurer* innerhalb der Redebegeleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Konstruktion in sechs Texten³⁰⁸ – vier davon stammen allein von Honoré de Balzac – mit zwischen einem und drei Tokens nachgewiesen wurde, ist sie bereits ab 1847 nicht mehr vorhanden und spielt damit im Vergleich zu *murmurer* als Voll-Redeverb nur eine untergeordnete Rolle.

(530) – *Donnez-vous donc mille peines pour établir convenablement des filles ! dit en murmurant monsieur de Navarreins au vidame.*³⁰⁹

(9951) – *Monsieur de Cagliostro ! murmura Jeanne chancelante et près de s'évanouir.*³¹⁰

Victor Hugo ist erneut der erste, der dieses expressive Redeverb einführt: Für *Han d'Islande* (1823) (Nr. 8) wurden innerhalb des redebegleitenden *discours citant* 28 konjugierte Formen von *murmurer* ermittelt. Innerhalb des Zeitraums von 1830 bis 1842 kommt *murmurer* nur in zwei Romanen als Redeverb vor: in *Notre-Dame de Paris* (1831) (Nr. 11) von Victor Hugo mit zehn Tokens und in *Consuelo* (1842) (Nr. 19) von George Sand mit acht Tokens. Danach wird es häufiger eingesetzt, auch wenn die Verteilung

³⁰⁸ Honoré de Balzac, *La peau de chagrin* (1831)(2 Tokens), *La femme de trente ans* (1842)(3 Tokens), *Le lys dans la vallée* (1844)(2 Tokens) und *Cousin Pons* (1847)(1 Token) und Victor Hugo, *Notre-Dame de Paris* (1831)(1 Token) und Eugène Sue, *Les mystères de Paris* (1842)(1 Token).

³⁰⁹ Honoré de Balzac, *La duchesse de Langeais* (1835).

³¹⁰ Alexandre Dumas, *Le collier de la reine* (1849).

sehr unregelmäßig ist. In 15 Romanen kommen jeweils zwischen 20 und 40 Tokens vor, für sieben Werke beträgt die Anzahl zwischen 50 und 100, und für drei Texte wurden 100 Tokens und mehr³¹¹ ermittelt. Ab 1900 verliert *murmurer* mit deutlich weniger als zehn Tokens pro Roman zunehmend an Bedeutung.

Interrompre kommt in 52 Texten innerhalb des redebegleitenden *discours citant* vor, davon in 32 Romanen als *verbum communicandi* und in 27 als Teil einer das Redeverb erweiternden *gérondif*-Konstruktion, z.B.

(453) – *Je vois, lui répondit sa femme en l'interrompant au milieu d'une tirade, que tu t'es endetté de deux cent mille francs ?*³¹²

(384) – *Ah ! ne l'écoutez pas, madame Bovary, interrompit Homais en se penchant sur son assiette, c'est modestie pure.*³¹³

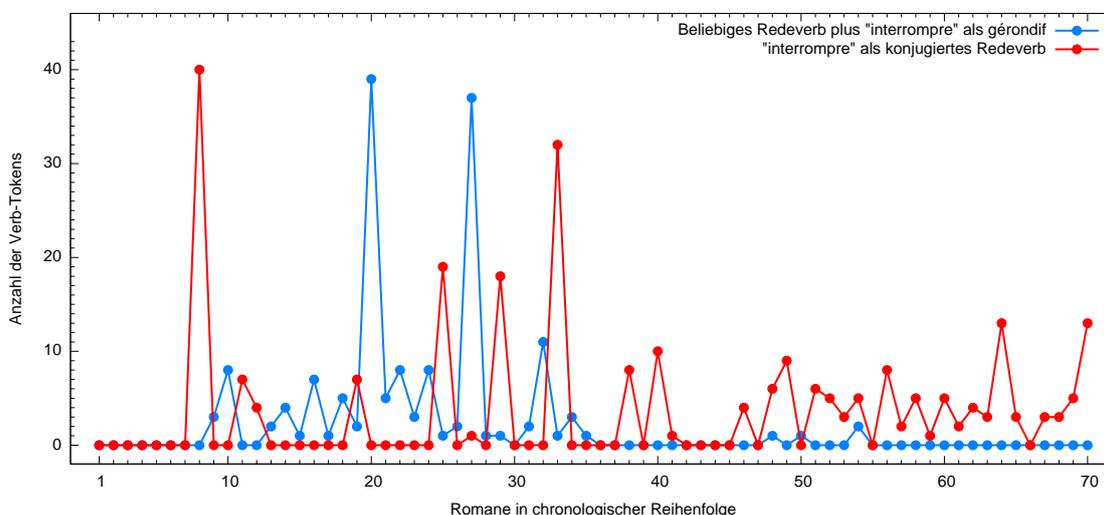


Abbildung 4.39: Entwicklung der konjugierten sowie der *gérondif*-Formen von *interrompre* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Bei Voltaire und Bernardin de Saint Pierre tritt *interrompre* in der Redebegleitung noch nicht in Erscheinung. Die *gérondif*-Konstruktion ist zwischen 1830 und 1848 in elf Texten mit fünf oder mehr Vorkommen präsent, von denen die meisten in *Les mystères de Paris* (1842) bzw. *Le juif errant* (1845) (Nr. 27) von Eugène Sue ermittelt wurden (42 bzw. 40 Beispiele). In *Cousine Bette* (1848) (Nr. 32) von Honoré de Balzac liegt die Anzahl von *en interrompant* letztmalig über zehn Tokens. Danach geht die Verwendung der *gérondif*-Form deutlich zurück, so dass bis zum Ende des Untersuchungszeitraums nur noch in

³¹¹ Alexandre Dumas, *Le comte de monte Christo* (1846)(138 Tokens), Alexandre Dumas, *Le collier de la reine* (1849)(111 Tokens) sowie Émile Zola, *Nana* (1879)(100 Tokens).

³¹² Honoré de Balzac, *César Birotteau* (1837).

³¹³ Gustave Flaubert, *Madame Bovary* (1857).

sechs Romanen zwischen einer und drei solcher Konstruktionen vorhanden sind. Nach 1879 kommt *interrompre* gar nicht mehr als *gérondif*-Form innerhalb der Redebegleitung vor. Als konjugiertes Redeverb fällt *interrompre* bis 1844 kaum auf – mit einer Ausnahme: Schon 1823 setzt Victor Hugo es in *Han d'Islande* gleich 40 Mal als Voll-Redeverb ein und legt damit das Maximum fest, das auch in später publizierten Romanen anderer Autoren nicht mehr erreicht wird. Alexandre Dumas macht von *interrompre* als *verbum communicandi* mit 19 Vorkommen in *Les trois mousquetaires* (1844) (Nr. 25), 18 in *Le comte de Monte Christo* (1846) (Nr. 29) und 33 in *Le collier de la reine* (1849) (Nr. 33) ebenfalls häufig Gebrauch. Auch nach 1850 ist es vielfach als Redeverb vertreten, wenn auch die großen Abweichungen nach oben hin fehlen und die maximale Anzahl von 13 Tokens nicht mehr übertroffen wird. Zwischen 1857 und 1922 setzen die Autoren *interrompre* in 23 von 32 Romanen mit zwischen einem und 13 Tokens ein, in 13 Texten liegt diese Zahl sogar zwischen fünf und 13³¹⁴.

Soupirer (Abb. 4.40) spielt als *verbum communicandi*, egal ob als konjugiertes Redeverb oder Teil einer *gérondif*-Paraphrase, eine nur marginale Rolle und ist lediglich in 20 Romanen in der Redebegleitung vorhanden. Der Gebrauch der *gérondif*-Konstruktion beschränkt sich auf den Zeitraum 1831 bis 1862, d.h. die Texte 10-39. In fünf Texten ist je eine *gérondif*-Konstruktion vorhanden, davon in zwei Romanen ausschließlich diese³¹⁵. Wieder ist Victor Hugo der erste, der *soupirer* als Voll-Redeverb einsetzt, jedoch stehen in *Notre-Dame de Paris* (1831) (Nr. 11) zwei *gérondif*-Formen ebenso viele konjugierte Formen gegenüber:

(114) – *Sont-ils heureux de voir tout cela, disait en soupirant Joannes de Molendino, toujours perché dans les feuillages de son chapiteau.*

(1412) – *Elle est peut-être sourde, dit Oudarde en soupirant.*

(2319) – *Eheu ! soupira Gringoire navré, bassa latinitas !*

(2393) – *Et moi ! soupira-t-elle faiblement.*³¹⁶

Fünf konjugierte Tokens oder mehr – das Maximum liegt bei acht Tokens – finden sich nur in vier Texten³¹⁷. Weitere elf Werke weisen nur jeweils ein oder zwei Verbtokens auf. Insgesamt ist die Verteilung zwischen 1831 und 1920 sehr lückenhaft, nur gegen Mitte des 19. und zu Beginn des 20. Jahrhunderts ist vereinzelt eine geringe Zunahme zu beobachten.

³¹⁴ Das Maximum für diesen Zeitraum, d.h. 13 Tokens, wird in *Les demi-vierges* (1894) (Nr. 64) von Marcel Prévost und *Le côté de Guermantes* (1922) (Nr. 70) erreicht.

³¹⁵ George Sand, *Consuelo* (1842) und Eugène Sue, *Le juif errant* (1845), mit jeweils einem Token.

³¹⁶ Alle vier Beispiele stammen aus *Notre-Dame de Paris* (1831) von Victor Hugo.

³¹⁷ Alexandre Dumas, *Le collier de la reine* (1849)(8 Tokens), Gustave Flaubert, *Madame Bovary* (1857)(5 Tokens), Marcel Proust, *Un amour de Swann* (1913)(5 Tokens) und Colette, *Chéri* (1920)(6 Tokens).

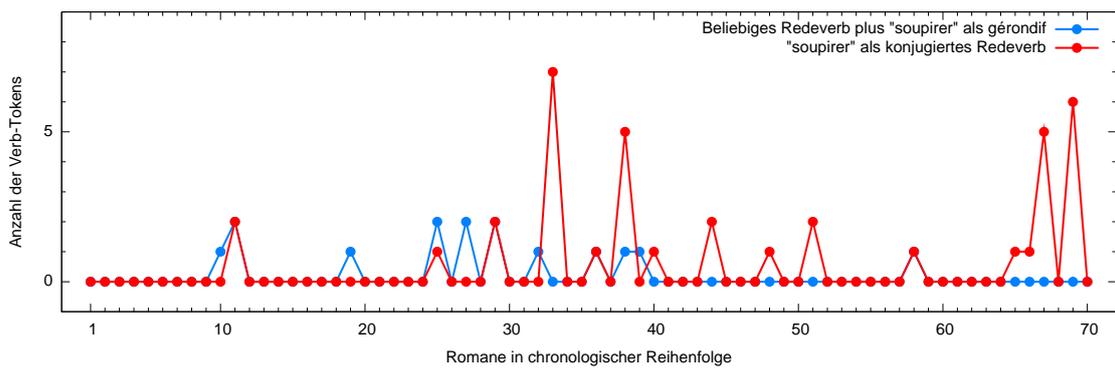


Abbildung 4.40: Entwicklung der konjugierten sowie der *gérondif*-Formen von *soupirer* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

Wie *soupirer* gehört auch ***éclater*** (Abb. 4.41) zu den seltenen Rede-Verben. Anders als die bisher besprochenen Verben steht es jedoch fast ausnahmslos in Verbindung mit einer *gérondif*-Konstruktion. *Éclater* wurde in 14 Romanen ermittelt, davon in zwölf als *gérondif*-Form in Kombination mit einem anderen Rede-Verb. Gebräuchlich sind vor

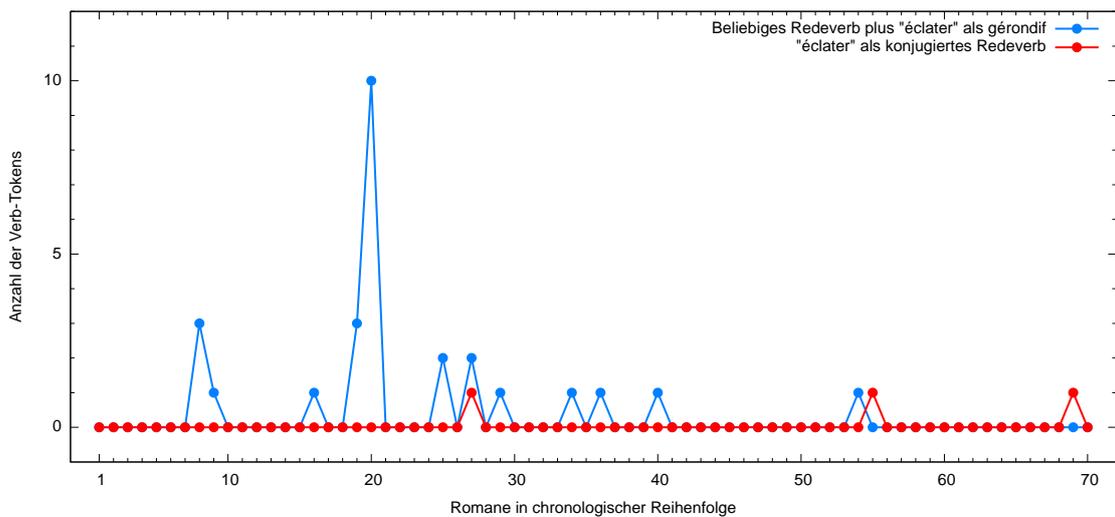


Abbildung 4.41: Entwicklung der konjugierten sowie der *gérondif*-Formen von *éclater* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

allem Wendungen wie *en éclatant de rire*, *en éclatant en sanglots* und *en éclatant* ohne Ergänzung, die im gesamten Korpus jeweils 23 Mal, zweimal bzw. einmal nachgewiesen wurden, z.B.:

(3249) – *Quel nom ! dit le prince en éclatant de rire ; mais pardon, il est sublime pour vous.*³¹⁸

(4184) – *Oh ! ne dites pas de pareilles choses ! s'écria la procureuse en éclatant en sanglots.*³¹⁹

(6524) – *Mais non, il voulait me faire emprisonner ce matin ! reprit Morel en éclatant, ce n'est donc pas lui ?... Oh ! tant mieux !... tant mieux !... Elle n'a pas même d'excuse à sa faute, je ne serai pour rien dans son déshonneur... Je pourrai sans remords la maudire !...*³²⁰

In *Le juif errant* (1845) (Nr. 27) von Eugène Sue steht einer *gérondif*-Form eine konjugierte Verbform von *éclater* gegenüber. Bei näherem Betrachten der Fundstelle³²¹ zeigt sich jedoch, dass es sich in diesem Fall um einen Fehltreffer handelt. Aus der Sicht der Software ist dies ein völlig korrektes Ergebnis, erfüllt es doch die für einen redebegleitenden Satz gestellte Bedingung, die die Abfolge Satzzeichen - Verb - nachgestelltes Subjekt verlangt (vgl. Abb. 3.4 bis 3.6). Dennoch ist in diesem Fall *éclater* kein *verbum dicendi*, sondern die direkte Rede wird mit einer Tokenabfolge weitergeführt, die zufällig der Redebegleitung entspricht. Genauso verhält es sich mit der Fundstelle aus *Bouvard et Pécuchet* (1881) (Nr. 55), wo ebenfalls eine vermeintliche konjugierte Form³²² ermittelt wurde. Der einzige korrekte Nachweis über die Verwendung einer konjugierten Form von *éclater* stammt damit aus *Chéri* (1920) (Nr. 69) von Colette:

(367) – *Tout de même, éclata-t-elle... elle t'a traité de maquereau !*³²³

Nach 1862 wird *éclater* als Teil einer *gérondif*-Konstruktion mit Ausnahme eines einzigen Nachweises in *Nana* (1879) (Nr. 54) von Émile Zola nicht mehr verwendet. Das einzige Beispiel mit *éclater* als konjugiertem *verbum dicendi* von Colette ist nur ein schwacher Hinweis darauf, dass das Verb ab diesem Zeitpunkt nur noch als konjugiertes Redeverb eingesetzt wird. Um dies nachzuweisen, müssten weitere, deutlich jüngere Texte analysiert werden, was allerdings aufgrund urheberrechtlicher Beschränkungen für diese Untersuchung nicht möglich ist.

³¹⁸ Stendhal, *Le rouge et le noir* (1830).

³¹⁹ Alexandre Dumas, *Les trois mousquetaires* (1844).

³²⁰ Eugène Sue, *Les mystères de Paris* (1842).

³²¹ Eugène Sue, *Le juif errant* (1845): (12542) – *Oui, reprit Rodin d'une voix solennelle, il avait des visions effrayantes... Cette jeune fille, morte pour lui en état de péché mortel, il la voyait plongée au milieu des flammes éternelles ! Sur son beau visage, défiguré par les tortures infernales, éclatait le rire désespéré des damnés...*

³²² Gustave Flaubert, *Bouvard et Pécuchet* (1881): (775) – *Mille excuses ! témoin celui de Lisbonne ! Quant à nos pays, les mines de houille et de pyrite martiale y sont nombreuses et peuvent très bien en se décomposant, former les bouches volcaniques. Les volcans, d'ailleurs, éclatent toujours près de la mer.*

³²³ Colette, *Chéri* (1920).

Interessant ist auch die Verteilung des Verbs *ricaner* (Abb. 4.42), das in 19 Romanen innerhalb der Redebegleitung nachgewiesen wurde. Zwar ist es keineswegs häufig, da

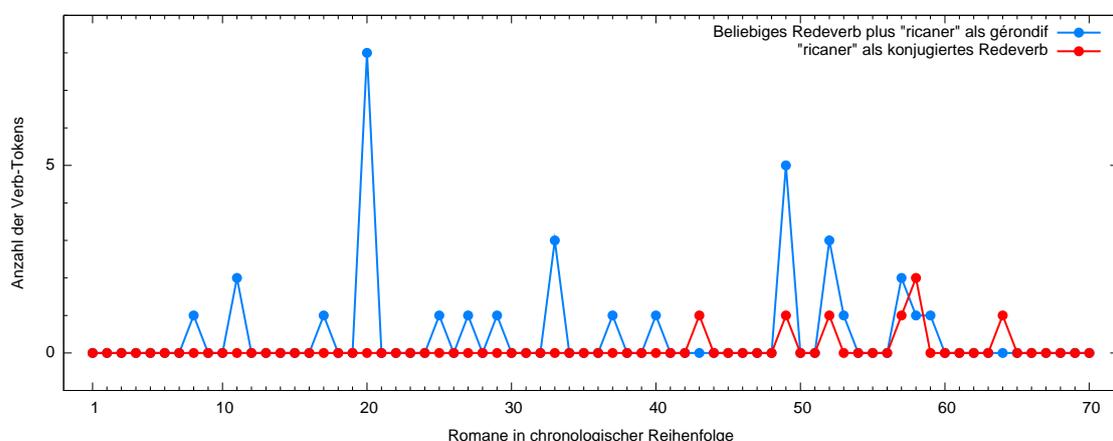


Abbildung 4.42: Entwicklung der konjugierten sowie der *gérondif*-Formen von *ricaner* innerhalb der Redebegleitung. Die Aufschlüsselung der Romannummern entlang der horizontalen Achse erfolgt über Tabelle 3.1.

15 Romane nur je ein oder zwei Tokens enthalten, dennoch gehört es zu den wenigen Verben, die hauptsächlich als *gérondif* in Verbindung mit einem *verbum dicendi* verwendet werden. Der erste Nachweis wurde 1823 bei Victor Hugo ermittelt, und nach 1839 setzen es auch andere Autoren ein:

(1411) – *Ne craignez rien de la part de Dieu, dit Hacket en ricanant.*³²⁴

Erst ab 1862 kommt es auch als konjugiertes Redeverb vor, und zwar in *Thérèse Raquin* (1867) (Nr. 43) von Émile Zola:

(887) – *Il était bon, oui, je sais, ricanait Laurent, tu veux dire qu'il était bête, n'est-ce pas....*³²⁵

Zola scheint eine besondere Vorliebe für das Verb zu haben, da er es als Einziger in sechs seiner Werke zwischen 1874 und 1890 innerhalb der Redebegleitung benutzt. Bis auf einen Roman überwiegt dabei die *gérondif*-Konstruktion. Nur in *La terre* (1887) sind die konjugierten Verbformen mit zwei Tokens in der Überzahl.

Für die vorgestellten Verben kann die eingangs aufgestellte These, dass vor allem *verba communicandi*, die zunächst nur Teil einer *gérondif*-Paraphrase sind, allmählich selbst zum vollwertigen Redeverb werden, nicht bestätigt werden. Eine solche Entwicklung trifft allenfalls auf das Verb *interrompre* zu, wo in der ersten Korpuschälfte neben einzelnen konjugierten Formen viele *gérondif*-Konstruktionen vorkommen und es um

³²⁴ Victor Hugo, *Han d'Islande* (1823).

³²⁵ Émile Zola, *Thérèse Raquin* (1876).

ca. 1850 einen Umkehrpunkt gibt, ab welchem die konjugierten Formen vorherrschen. In der Regel ist jedoch die konjugierte Form weit in der Überzahl, und zwar bereits zu dem Zeitpunkt, wo parallel dazu noch *gérondif*-Paraphrasen gebraucht werden (*balbutier* und *murmurer*). Außerdem gibt es einzelne expressive Verben, z.B. *éclater* und *ricaner*, die vor allem in *gérondif*-Paraphrasen stehen, aber in kaum signifikanter Zahl auch als konjugierte Voll-Redeverben gebraucht werden. Andere Verben tauchen – oft zum Ende des Untersuchungszeitraums hin – als konjugiertes Redeverb auf, ohne dass es vorher *gérondif*-Konstruktionen in nennenswerter Zahl gegeben hätte (z.B. *bégayer*, *crier*, *s'écrier*, *hurler* und ggf. *soupirer*). Insgesamt ist die *gérondif*-Konstruktion im Vergleich zu der konjugierten Form zu selten, als dass man sie als Auslöser für eine so weitreichende Veränderung sehen könnte. Fest steht, dass die expressiven Verben ab der zweiten Korpushälfte zunehmen und gleichzeitig die Dominanz von *dire* als *Passepartout*-Redeverb abnimmt, was aber in erster Linie auf neue literarische Konventionen sowie idiolektale Vorlieben der Schriftsteller zurückzuführen sein dürfte.

Die übrigen von Vivienne Mylne (1987) genannten und hier nicht besprochenen Verben können die These des Wandels von *gérondif*- zu konjugierter Form aufgrund ihrer sehr geringen Frequenz, z.B. kommen *chuchoter* und *vociférer* nur in acht bzw. fünf Romanen mit nur je einem Token vor, ebenfalls nicht stützen. Auch für *souffler* und *glapir* lässt sich nicht von einer Entwicklung von der *gérondif*-Konstruktion hin zum Voll-Redeverb sprechen, da sie beide ohnehin nur als konjugierte Verbform nachgewiesen wurden. *Nasiller* wiederum wurde nur in einem einzigen Text in Gestalt eines *gérondif*-Tokens gefunden.

4.2.6 *dire* und *parler* als Teil der *gérondif*-Paraphrase

Das Redeverb *dire*, das in fast allen untersuchten Romanen das am häufigsten gebrauchte *verbum communicandi* ist³²⁶, wird vereinzelt auch innerhalb der *gérondif*-Konstruktionen gebraucht. Gleiches gilt für das Verb *parler*, das jedoch im Gegensatz zu *dire* kaum als Vollredeverb in Erscheinung tritt. Aufgrund ihrer Verteilung lohnt sich ein Blick auf die beiden Konstruktionen *en disant* und *en parlant* innerhalb des redeeinleitenden und redebegleitenden *discours citant*.

Im gesamten Arbeitskorpus sind insgesamt lediglich 31 *gérondif*-Paraphrasen mit

³²⁶ Innerhalb der Redebegleitung ist *dire* mit Ausnahme der vier untersuchten Romane von Jules Verne mit Abstand das häufigste Redeverb (vgl. Abb. 4.10; *répondre* ist hier deutlich frequenter als *dire*). Abbildung 4.13 zeigt, dass in *Les demi-vierges* (1894) von Marcel Prévost *dire* und *faire* gleich aufliegen, doch ein Blick in Tabelle A.4 zeigt, dass *dire* auch dort, wenn auch nur knapp mit einem Token, überwiegt.

en disant enthalten³²⁷, die ausnahmslos als Teil eines redeenleitenden *discours citant* identifiziert wurden.

En disant bildet in fast allen Fundstellen die Erweiterung eines Nicht-Redeverbs und steht unmittelbar vor der direkten Rede, die nach einem abschließenden Satzzeichen – einem Doppelpunkt – beginnt. Bei Émile Zola und Marcel Proust wurde außerdem je ein Beispielpaar ermittelt, in denen die Redeverben *interrompre* und *conclure* innerhalb der Redeeinleitung mit *en disant* ergänzt werden, während innerhalb der Redebegleitung keine Ergänzung notwendig zu sein scheint:

(225) *Saccard étant venu s'asseoir à quelques pas des jeunes gens, M. de Mussy garda le silence et Maxime conclut en disant* : – *Moi, si j'étais à votre place, j'agis très cavalièrement.*³²⁸

(1462) – *Je te dis tout cela, conclut le jeune homme, pour que tu n'aies pas l'air d'une sottise....*³²⁹

(230) *Cela ne me parut pas certain ; mais cela dut me paraître probable, car je me sentis devenir très rouge, quand mon père m'interrompit en disant* : – *Je connais ces maisons-là ; j'en ai vu une, elles sont toutes pareilles ; Swann occupe simplement plusieurs étages, c'est Berlier qui les a construites.*³³⁰

(495) – *Oh ! mon ami, interrompit maman, rien ne prouve que ce soit vrai. On dit tant de choses.*³³¹

Möglicherweise wäre der Übergang zur direkten Rede ohne die *gérondif*-Form nur in Form eines die Figurenrede ankündigenden graphischen Markers zu abrupt.

Die *gérondif*-Konstruktionen, die *en disant* enthalten, lassen sich jeweils in verschiedene Gruppen unterteilen:

en disant + Ø

Gut ein Drittel dieser Konstruktionen kommt ohne weitere Ergänzung aus und besteht nur aus *en disant*, an das sich die direkte Rede nahtlos anschließt, z.B.

(3605) *Elle mit la pièce dans sa poche, et se borna à jeter un regard farouche à l'enfant en disant* : – *Que cela ne t'arrive plus, toujours !*³³²

³²⁷ Grundlage für die Suche sind alle Turns, die eine Redeeinleitung oder -begleitung enthalten. Nicht berücksichtigt werden Turns, in deren Einleitung oder Begleitung zwar *en disant* enthalten ist, das jedoch seinerseits eine Erweiterung einer Erweiterung ist, z.B. (327) – *Hé ! bien, quoi ? s'écria Grandet au moment où Cruchot lui mettait le journal sous les yeux en lui disant*: ... (Honoré de Balzac, *Eugénie Grandet* (1843)).

³²⁸ Émile Zola, *La curée* (1871).

³²⁹ Émile Zola, *La curée* (1871).

³³⁰ Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918).

³³¹ Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918).

³³² Victor Hugo, *Les misérables* (1862).

en disant + Dativobjekt

Weitere neun *gérondif*-Paraphrasen referenzieren mittels einer Dativ-Erweiterung auf einen Gesprächspartner, z.B.

(1606) *Mais le patron m'en refusa une autre en me disant* : – Non, monsieur, je ne peux pas gêner tout le monde pour vous.³³³

en disant + Adverbialergänzung

Zu einer weiteren Gruppe können diejenigen vier *gérondif*-Konstruktionen zusammengefasst werden, die adverbial erweitert werden, z.B.

(953) *À la vue de l'hôte et des gens de l'auberge, Morok se releva en disant d'une voix solennelle* : – Merci, mon Dieu ! d'avoir pu vaincre encore une fois par la force que vous m'avez donnée.³³⁴

en disant + Dativobjekt + Adverbialergänzung

In weiteren fünf Beispielsätzen wird die Konstruktion *en disant* sowohl durch Bezugnahme auf einen Gesprächspartner als auch durch eine adverbiale Ergänzung näher bestimmt, wie z.B. in

(329) *Il congédia son frère en lui disant avec un sourire* : – Oui, un nom à aller au bain ou à gagner des millions.³³⁵

Verglichen mit der Vielzahl der im Arbeitskorpus enthaltenen *gérondif*-Konstruktionen in redeeinleitenden, vor allem aber in redebegleitenden Sätzen, stellen die Konstruktionen mit *en disant* jedoch nur eine Randerscheinung dar.

Auch *parler* wird innerhalb der *gérondif*-Konstruktionen verwendet. Im vollständigen Arbeitskorpus wurden für die Redeeinleitung und Redebegleitung insgesamt 49 Beispiele dafür ermittelt. Die Verteilung ist hier allerdings eine ganz andere als bei *dire*: *en parlant* ist mit 42 Vorkommen vor allem in der Redebegleitung vertreten, während auf die Redeeinleitung nur sieben entfallen, z.B.

(6625) *Bientôt il dit à voix basse et comme en se parlant à lui-même* : (6626) – Est-il possible, mon Dieu ! que le hasard produise de pareilles ressemblances !³³⁶

Anders als bei *en disant* modifiziert die *gérondif*-Konstruktion hier immer ein „echtes“ *verbum communicandi*. Innerhalb der redeeinleitenden Sätze ist dies ausnahmslos *dire*. In den redebegleitenden Sätzen geht den *en parlant*-Konstruktionen in 23 Fällen *dire* voraus, gefolgt von *ajouter* und *repandre* (je fünf Vorkommen), *continuer* (drei Vorkom-

³³³ Marcel Proust, *Le côté de Guermantes* (1922).

³³⁴ Eugène Sue, *Le juif errant* (1845).

³³⁵ Émile Zola, *La curée* (1871).

³³⁶ Eugène Sue, *Le juif errant* (1845).

men), *répéter* (zwei Vorkommen) sowie *crier*, *faire*, *murmurer* und *répondre* (je ein Vorkommen). Die Paraphrasen können in drei Gruppen unterteilt werden:

en parlant + Adverbialergänzung

In 18 Beispielsätzen wird *en parlant* jeweils adverbial erweitert, wie z.B. in

(332) – *C'est tout ce que j'en ai, dit la marchande en parlant comme une flûte enrouée*.³³⁷

en se parlant à lui-même/elle-même

Eine weitere, fast ebenso große Gruppe bilden 17 redeenleitende und redebegleitende Sätze, die ein Gespräch der Figur mit sich selbst abbilden, konstruiert mithilfe von *en se parlant* in Kombination mit *à lui-même/elle-même*³³⁸:

(2675) – *Mon Dieu, mon Dieu! disait Haydn en se parlant à lui-même lorsqu'elle eut fini, je n'avais jamais entendu chanter; je ne savais pas ce que c'est que le chant!*³³⁹

en parlant + de

In 12 Fällen wird *en parlant* durch die Präposition *de* ergänzt, z.B.

(178) – *Ils font cuire trop à la va-vite, répondit-elle en parlant des grands restaurateurs, et puis pas tout ensemble*.³⁴⁰

was für den Schriftsteller im Bedarfsfall ein probates Mittel zur Verdichtung des Dialogs darstellt, indem zwar auf den Gesprächsgegenstand referenziert wird, Details aber nicht in Form direkter Rede wiedergegeben werden.

In einem weiteren Beispiel wird *en se parlant* mit einer adverbialen Ergänzung kombiniert, bei der *à lui-même* elidiert wird:

(2979) – *Eh bien! elle ne m'aime plus, se répétait-il en se parlant tout haut comme pour s'apprendre sa position*.³⁴¹

Während die Redebegleitung im vorangegangenen Beispiel bereits etwas sperrig anmutet – für Zola wäre zu einem späteren (Schaffens-)Zeitpunkt sicher auch eine Elision des *en se parlant* denkbar gewesen – so wirkt der *discours citant* in folgendem Beispiel von Stendhal durch das Aneinanderreihen von VC plus sowohl *gérondif*-Konstruktion als auch Partizipialkonstruktion erst recht schwerfällig, ganz zu schweigen davon, dass die Wahl von *en parlant* den Kontext nicht unbedingt adäquat wiedergibt:

³³⁷ Honoré de Balzac, *César Birotteau* (1837).

³³⁸ Ein Beispiel in *Le juif errant* (1845) von Eugène Sue lautet allerdings: (9567) *Au lieu de répondre, M. de Montbron semblait de plus en plus absorbé, pensif, en contemplant la jeune fille, et il ne put s'empêcher de dire en se parlant à soi-même: – Non... non... c'est impossible... et pourtant...*

³³⁹ George Sand, *Consuelo* (1842).

³⁴⁰ Marcel Proust, *À l'ombre des jeunes filles en fleurs* (1918).

³⁴¹ Stendhal, *Le rouge et le noir* (1830).

(491) *Pendant cette délibération qui fut longue, la comtesse avait lié conversation avec la jeune fille qui était à pied sur la route et dans la poussière à côté de la calèche ; elle avait été frappée de sa beauté. – Le soleil va vous faire mal, mademoiselle ; ce brave soldat, ajouta-t-elle en parlant au gendarme placé à la tête des chevaux, vous permettra bien de monter en calèche.*³⁴²

Dass *en disant* ausschließlich innerhalb der Redeeinleitung zu finden, *en parlant* hingegen weitgehend auf die Verwendung innerhalb der Redebegleitung beschränkt ist, lässt sich zum einen anhand der unterschiedlichen Grundbedeutungen der beiden Verben erklären. Die Verwendung des Universalredeverbs *dire* ist laut Definition³⁴³ mit einer Kommunikationsabsicht verbunden, während *parler* sich in seiner Grundbedeutung³⁴⁴ ganz allgemein auf die Artikulation von Lauten bezieht. Für die Kombination aus Redeverb und *gérondif*-Ergänzung gilt außerdem, dass ein „semantisches Inklusionsverhältnis“³⁴⁵ vorliegen muss. Ist *dire* nun Teil der *gérondif*-Konstruktion, so darf das ihm vorausgehende Redeverb streng genommen kein *verbum communicandi* sein, weil beide Verben sonst auf derselben semantischen Hierarchieebene liegen würden. Betrachtet man die in diesem Kapitel genannten Beispiele, so ist diese Bedingung für *en disant* in der Regel erfüllt, da die *gérondif*-Form vor allem Nicht-Redeverben erweitert: *jeter (un regard à quelqu'un) en disant* (S. 173), *refuser (quelque chose à quelqu'un) en disant* (S. 174), *se relever en disant* und *congédier en disant* (S. 174). Bei den beiden Beispielen *conclure en disant* (S. 173) und *interrompre quelqu'un en disant* (S. 173) handelt es sich um Sonderfälle, da beide Verben keine *verba communicandi* im eigentlichen Sinne sind³⁴⁶, obwohl sie vereinzelt auch als Voll-Redeverben gebraucht werden. In allen aufgeführten Beispielen wird die *gérondif*-Konstruktion *en parlant* in der Redebegleitung durchgängig in Verbindung mit klassischen Redeverben – z.B. *ajouter, dire, répéter* und *répondre* – verwendet und dient dazu, diese näher zu bestimmen, d.h. gemäß der Grundbedeutung von *parler* Auskunft über die Art und Weise des Sprechens zu geben.

³⁴² Stendhal, *La chartreuse de Parme* (1839).

³⁴³ Definition von *dire*: Énoncer un propos par la parole physiquement articulée avec l'intention de le communiquer et d'appeler éventuellement une réponse ou une réaction. (Quelle: Le Trésor de la Langue Française Informatisé (<http://atilf.atilf.fr/>)).

³⁴⁴ Definition von *parler*: *Articuler les sons d'une langue naturelle*. (Quelle: Rey-Debove & Rey (1993), S. 1591.

³⁴⁵ Annette Sabban (1988), S. 52.

³⁴⁶ Definition von *interrompre*: *Arrêter quelqu'un dans son discours, lui couper la parole*. Definition von *conclure*: *Mettre un terme à quelque chose qui est en cours*. (Quelle für beide Angaben: Le Trésor de la Langue Française Informatisé (<http://atilf.atilf.fr/>)).

4.3 Vergleich mit bisherigen Untersuchungen

Auch wenn dem *discours citant* von wissenschaftlicher Seite bislang nur wenig Beachtung zuteil wurde, gibt es einzelne Literatur- und Sprachwissenschaftler, die den *discours citant* trotz des damit verbundenen «travail fastidieux»³⁴⁷ anhand ausgewählter literarischer Werke untersucht haben und unter anderem Ergebnisse liefern, die sich grundsätzlich zum Vergleich mit den von DCAnalyse gewonnenen Daten eignen sollten.

In Prince (1980a), Prince (1980b), Halina Suwala (1987) und Fairon (2004) werden jeweils konkrete Zahlen zu einzelnen Romanen, die auch im Arbeitskorpus enthalten sind, genannt. Die meisten dieser Ergebnisse unterscheiden sich jedoch stark von denen, die DCAnalyse liefert, denn die von der Software errechneten Werte liegen oft deutlich sowohl darüber als auch darunter. Im Folgenden werden die Ergebnisse einander gegenübergestellt und kommentiert.

Prince (1980a) setzt sich mit dem «discours attributif» in *La peste*³⁴⁸ (1947) von Albert Camus auseinander. Um die ermittelten Zahlen besser einordnen zu können, zieht Prince vereinzelt Werte zu anderen Romanen hinzu. So beziffert er die Menge der Redeverbren für *Eugénie Grandet* (1843) von Honoré de Balzac auf zwölf³⁴⁹. Der Leser erfährt jedoch nicht, ob sich diese Angabe auf die gesamte direkte Rede bezieht, d.h. für Redeeinleitung und Redebegleitung zusammengenommen oder nur auf eine von beiden. DCAnalyse hat hingegen 19 *verba communicandi* für die Redeeinleitung sowie 20 für die Redebegleitung ermittelt, woraus eine Vereinigungsmenge von 30 *verba communicandi* für die komplette direkte Rede resultiert. An anderer Stelle nimmt Prince Bezug auf den Anteil der redееinleitenden Sätze in *Madame Bovary* (1857) von Gustave Flaubert, der nach seiner Rechnung 25% ausmacht. DCAnalyse zufolge ist der Anteil der Redeeinleitung mit 12,4% gerade einmal halb so groß, während derjenige für die Redebegleitung 56,4% beträgt und insgesamt 31,2% der Turns unbegleitet sind.

In einem weiteren Artikel beschäftigt sich Prince (1980b) ausschließlich mit *Madame Bovary* von Gustave Flaubert. Abgesehen von grundsätzlichen Erläuterungen zum «discours attributif» und zur Bedeutung der darin vorkommenden Figurbezeichnungen führt Prince auch genaue Zahlen bezüglich der Verteilung des «discours attributif» bzw. der verwendeten Redeverbren im Roman an. So sind in *Madame Bovary* 1262³⁵⁰

³⁴⁷ Siehe Zitat auf Seite 11.

³⁴⁸ Dieser Roman konnte aus urheberrechtlichen Gründen nicht in das für diese Arbeit verwendete Korpus aufgenommen werden.

³⁴⁹ Prince (1980a), S. 101.

³⁵⁰ Prince (1980b), S. 271.

(1102)³⁵¹ Turns enthalten, von denen 879 (966) entweder eingeleitet oder begleitet werden. Innerhalb der direkten Rede beläuft sich die Zahl der *verba communicandi* auf insgesamt 44³⁵² (63). *Dire* hat mit 362 (275) Vorkommen einen Anteil von mehr als 30% (37,7%) an den Redeverben. Der Anteil der «propositions devançant» beträgt 25%³⁵³ (12,4%). Trotz Flauberts Streben nach lexikalischer Vielfalt³⁵⁴ kommen manche Verben laut Prince überhaupt nicht als Redeverb vor, wie z.B. *affirmer*, *remarquer*, *insister* und *admettre*³⁵⁵. Bis auf eine Ausnahme stimmt diese Aussage exakt mit den Ergebnissen von DCAnalyse überein: *insister* wurde immerhin ein einziges Mal in einem redееinleitenden *discours citant* gefunden:

(1699) *Alors Charles insista* : – *Tu reviendrais dimanche. Voyons, décide-toi ! tu as tort, si tu sens le moins du monde que cela te fait du bien.*³⁵⁶

Da Prince sich auch in dieser Untersuchung nicht zur Art und Weise der Datenerhebung äußert und die identifizierten Redeverben zwar auflistet, aber keine Angaben zu deren Frequenz macht, bleibt jeder Versuch, die Abweichungen zwischen den von ihm und den von DCAnalyse ermittelten Werten zu erklären, reine Spekulation. Allerdings kann man davon ausgehen, dass beim manuellen Erfassen der 1262 Dialogstellen – davon ist aufgrund des Publikationsjahrs des Artikels auszugehen – Fehler passieren.³⁵⁷ Ob diese allerdings so zahlreich sind, dass sie zu so deutlichen Unterschieden führen, kann jedoch auch nicht sicher behauptet werden.

Halina Suwala (1987) widmet sich dem *discours attributif* in *La curée* (1871) von Émile Zola und untersucht unter anderem, wie häufig sich Figuren an Gesprächen beteiligen und mit welchen Bezeichnungen sie innerhalb von Gesprächen versehen werden. Darüber hinaus liefert die Autorin eine Übersicht darüber, welche Redeverben wie oft enthalten sind und mit wie vielen verschiedenen Verben sich die Sprechenden zu Wort melden. Die den *discours citant* betreffenden Zahlen weichen jedoch erheblich von denen ab, die DCAnalyse ermittelt hat und in Tabelle 4.1 gegenübergestellt werden. Ein Vergleich der Ergebnisse ist nur bedingt möglich, da nicht erwähnt wird, wie Suwala die-

³⁵¹ Der von DCAnalyse ermittelte Wert wird jeweils in Klammern hinter den von Prince gesetzt.

³⁵² Prince (1980b), S. 273. Eine Liste dieser Verben – jedoch ohne Frequenzangaben – findet sich auf S. 275. Aufgeführt sind auch Redewendungen wie *faire des exclamations*, *laisser tomber (des mots)*, *prendre la défense de* und *ne pas retenir une phrase*. Dies erklärt jedoch nicht, wieso Prince insgesamt mehr Turns gefunden, aber gleichzeitig viel weniger Redeverben identifiziert hat.

³⁵³ Prince (1980a), S. 102.

³⁵⁴ Vgl. Prince (1980b), S. 273.

³⁵⁵ Vgl. Prince (1980b), S. 273.

³⁵⁶ Gustave Flaubert, *Madame Bovary* (1857).

³⁵⁷ So ist Prince's Aussage, *insister* sei nicht als Redeverb enthalten, schlicht falsch, was DCAnalyse durch das oben genannte Beispiel (1699) beweist.

	ZOL_CUREE		
	Suwala1987	DCAnalyse	
		RE	RB
Summe VC-Types	26	43	22
<i>dire</i>	179	23	119
<i>murmurer</i>	67	4	34
<i>répondre</i>	57	3	44
<i>demander</i>	41	5	24
<i>continuer</i>	39	8	19
<i>s'écrier</i>	27	5	14
<i>reprandre</i>	27	10	12
<i>ajouter</i>	18	2	1
<i>répéter</i>	17	4	7
<i>crier</i>	16	4	7
<i>balbutier</i>	10	2	4
<i>interrompre</i>	6	1	4
<i>conclure</i>	6	1	3
<i>penser</i>	5	1	4
<i>insister</i>	3	1	0
<i>poursuivre</i>	2	0	2
<i>répliquer</i>	2	0	2

Tabelle 4.1: Vergleich der Ergebnisse von Suwala (1987) (S. 56) und DCAnalyse für den Roman *La curée* (1871) von Émile Zola. RE = Redeeinleitung, RB = Redebegleitung.

se gewonnen hat.³⁵⁸ So beziffert die Autorin den Anteil der «scènes dialoguées»³⁵⁹ auf 34%, ohne jedoch genauer zu erläutern, wie sie diesen berechnet hat. Davon entfallen wiederum 35%³⁶⁰ auf die Redeeinleitung, während sich die Redebegleitung ebenfalls auf 35% beläuft. DCAnalyse hat für *La curée* völlig andere Werte ermittelt: Der Gesamtanteil der direkten Rede beträgt nur 12,1%, der sich seinerseits aus 24,5% Redeeinleitung, 54,4% Redebegleitung und 21% alleinstehenden Turns zusammensetzt. Diese Werte sind insofern plausibel, als sie sich nur unwesentlich von den für die anderen 15 Romane von Zola ermittelten unterscheiden.³⁶¹ Den Anteil des *discours citant*, der ein *participe présent* oder ein *gérondif* enthält, beziffert Suwala auf 9,5%, ohne die beiden Einzelwerte zu nennen. Das Ergebnis von DCAnalyse stimmt hiermit fast überein³⁶², jedoch

³⁵⁸ Aufgrund des Erscheinungsjahrs des Artikels kann man davon ausgehen, dass die Autorin den Text manuell bearbeitet hat. Man erfährt nicht, ob sich die Angaben zu den verwendeten Verben auf die Redeeinleitung oder die Redebegleitung bzw. beide beziehen, auch wenn grundsätzlich, d.h. eine Seite später, auf die Unterscheidung zwischen Redeeinleitung und Redebegleitung hingewiesen wird.

³⁵⁹ Halina Suwala (1987), S. 53.

³⁶⁰ Halina Suwala (1987), S. 57.

³⁶¹ Laut DCAnalyse liegt der Anteil der direkten Rede in den untersuchten Texten von Émile Zola zwischen 11,9% (*Thérèse Raquin* (1867)) und 41,1% (*La conquête de Plassans* (1874)). In zehn von 16 Romanen liegt der Anteil außerdem unter der 20%-Marke.

³⁶² Siehe Tabelle C.1: Die Addition von Typ 2a (Konstruktionen mit *participe présent*) und Typ 3b (Konstruktionen mit *gérondif*) ergibt 9,7%.

nur, wenn man davon ausgeht, dass es sich um den redebegleitenden *discours citant* handelt. Doch auch dieses nützliche Detail findet keine Erwähnung. Die Zahl der einfachen Redeverben, «verbes d'attribution simples»³⁶³, beläuft sich Suwala zufolge auf 26 (49)³⁶⁴. Obwohl sie insgesamt weit weniger Redeverben zählt, sind die einzelnen Verben im Vergleich zu den von DC-Analyse ermittelten sehr viel häufiger vertreten, und nur bei einzelnen Verben gibt es mehr oder weniger exakte Übereinstimmungen: *répondre* 57 (46), *reprendre* 27 (22), *crier* 16 (11), *balbutier* 10 (6), *conclure* 6 (4), *penser* 5 (5) und *insister* 3 (1) Vorkommen. Gerade bei den Verben, die auch Halina Suwala nur in geringem Umfang nachgewiesen hat, kommt es zu Übereinstimmungen, d.h. im Falle von *penser*, *poursuivre* und *répliquer*.

Eine Untersuchung neueren Datums, die außerdem computergestützt vorgegangen ist und Ergebnisse mit Hilfe der Software INTEX³⁶⁵ gewonnen hat, ist Fairon (2004). Zusätzlich zu zwei Presse-Korpora (*Le Monde* und *Le soir*) wurden auch vier literarische Texte untersucht: je zwei aus dem 19. Jahrhundert, *Le rouge et le noir* (1830) von Stendhal und *Notre-Dame de Paris* (1831) von Victor Hugo³⁶⁶ sowie zwei aus dem 20. Jahrhundert³⁶⁷. Mittels regulären Ausdrucks werden alle Konstruktionen herausgefiltert, deren Struktur dem redebegleitenden *discours citant* entspricht. Fehltreffer werden im Anschluss daran manuell aussortiert.³⁶⁸ Analysiert werden die Häufigkeit der redebegleitenden Konstruktionen («incises») und der darin auftretenden *verba communicandi*. In Tabelle 4.2 werden die Ergebnisse von Fairon (2004) (S. 199, Tabelle 4) und DC-Analyse zu den Romanen *Le rouge et le noir* (1830) von Stendhal und *Notre-Dame de Paris* (1831) von Victor Hugo gegenübergestellt. Während sich bezüglich der Anzahl der Fundstellen mit redebegleitendem *discours citant* doch deutliche Unterschiede zeigen, d.h. 1312 (Fairon (2004)) gegenüber 1652³⁶⁹ (DC-Analyse), ist die Summe der VC-Types für *Le rouge et le noir* (1830) jeweils identisch, und auch für *Notre-Dame de Paris* (1831) wurden annähernd gleich viele VC-Types ermittelt, d.h. 31 (Fairon (2004)) bzw. 34 (DC-

³⁶³ Halina Suwala (1987), S. 56.

³⁶⁴ Der von DC-Analyse ermittelte Wert wird jeweils in Klammern hinter den von Suwala gesetzt.

³⁶⁵ <http://intex.univ-fcomte.fr/>.

³⁶⁶ Fairon (2004), S. 196.

³⁶⁷ Dabei handelt es sich um die Romane *Oubliez Adam Weinberger* (1999) von Vincent Engel und *Jusqu'à preuve du contraire* (1995) von Bérengère Deprez.

³⁶⁸ Vgl. Fairon (2004), S. 197. Auch wenn die Korpusgröße noch überschaubar ist, so stellt das manuelle Aussortieren auch hier einen nicht zu unterschätzenden Aufwand dar, der bei noch größeren Korpora sicherlich nicht mehr vertretbar ist. Interessant wäre zudem zu wissen, wieviele der Fundstellen tatsächlich unbrauchbar waren.

³⁶⁹ Vgl. Fairon (2004), S. 197. Die geringere Anzahl von Dialogstellen erklärt sich zumindest teilweise durch manuelles Aussortieren nach dem Ende der Suche. Diese Methode scheidet für die vorliegende Arbeit als Korrektiv angesichts der beträchtlichen Korpusgröße aus.

	STE_ROUGEN		HUG_NODAME	
	Fairon2004	DCAnalyse	Fairon2004	DCAnalyse
Anzahl Einschübe	1312	1652	1248	1509
Summe VC-Types	17	17	31	34
<i>demander</i>	0	0	45	38
<i>grommeler</i>	0	0	18	13
<i>murmurer</i>	0	0	22	10
<i>observer</i>	0	0	15	14
<i>poursuivre</i>	0	0	27	22
<i>repartir</i>	0	n. vorh.	15	n. vorh.
<i>s'écrier</i>	0	72	64	48
<i>crier</i>	7	6	91	60
<i>répliquer</i>	9	7	13	11
<i>répondre</i>	13	56	116	113
<i>répéter</i>	14	13	23	10
<i>continuer</i>	34	26	10	10
<i>reprendre</i>	64	48	148	134
<i>ajouter</i>	91	63	29	23
<i>penser</i>	162	53	18	10
<i>se dire</i>	273	102	11	3
<i>dire</i>	578	433	547	415

Tabelle 4.2: Vergleich der Ergebnisse von Fairon (2004) (S. 199) und DCAnalyse zu den Romanen *Le rouge et le noir* (1830) von Stendhal und *Notre-Dame de Paris* (1831) von Victor Hugo.

Analyse). Vergleicht man nun die Vorkommen der einzelnen *verba communicandi* innerhalb der Redebegleitung, so unterscheiden sich die Zahlen deutlich voneinander. Für *Le rouge et le noir* (1830) stimmen die von DCAnalyse ermittelten Vorkommen nur für die Verben *crier*, *répliquer* und *répéter* fast überein, für alle anderen in Tabelle 4.2 aufgeführten ergeben sich signifikante Abweichungen. Gleiches gilt für *Notre-Dame de Paris* (1831): Nur für *observer*, *répliquer* und *répondre* stimmen die Werte annähernd überein, und nur *continuer* wurde gleich oft ermittelt. Die großen Unterschiede in der Anzahl der ermittelten redebegleitenden Sätze führen zwangsläufig zu voneinander abweichenden Zahlen, jedoch erklären sie nicht, weshalb die Werte für die einen Verben fast übereinstimmen, für andere aber nach oben oder unten erheblich abweichen. Nicht nachvollziehbar ist außerdem, weshalb Fairon (2004) für *Le rouge et le noir* (1830) überhaupt keine Redebegleitung mit *s'écrier* nachgewiesen hat, obwohl DCAnalyse 72 Vorkommen gezählt hat, wie z.B.

(2637) – *Tout doit être singulier dans le sort d'une fille comme moi, s'écria Mathilde impatientée.*

(3640) – *Non, je ne le suis pas pour vous, s'écria Julien en tombant à ses genoux.*³⁷⁰

Die Gegenüberstellung der Ergebnisse aus den vier Untersuchungen mit denen von DCAnalyse zeigt, dass es zum Teil erhebliche Unterschiede gibt. Deren Zustandekommen zu erklären bzw. die „Stimmigkeit“ der einzelnen Werte einzuschätzen, ist vor allem für die drei weiter zurückliegenden, aus prä-informatischer Zeit stammenden Arbeiten schwierig, da man kaum etwas über die Methoden der Datenerhebung erfährt. Warum in Prince (1980a), Prince (1980b) und Halina Suwala (1987) die gefundene Anzahl von *verba communicandi* jeweils deutlich unter der von DCAnalyse liegt, ist schlicht nicht nachvollziehbar. Es ist jedoch anzunehmen, dass beim manuellen Exzerpieren von Textstellen trotz größter Sorgfalt Fehler entstehen, indem Fundstellen entweder gar nicht erst entdeckt oder bei der späteren Zählung übersehen werden. Dies zumindest trifft auf einen durch den Computer analysierten Text nicht zu, auch wenn sich Fehler ganz anderer Art ergeben, weil dieser entweder alles oder nichts findet.

³⁷⁰ Beide Beispiele stammen aus *Le rouge et le noir* (1830) von Stendhal.

Kapitel 5

Fazit und Ausblick

Die vorliegende Arbeit hat sich mit dem redebegleitenden *discours citant*, wie er in enger Verbindung mit der Figurenrede häufig in Prosatexten anzutreffen ist, auseinandergesetzt und diesen nach verschiedenen Gesichtspunkten und unter Einbeziehung einer Vielzahl von Texten, also korpusbasiert, untersucht. Die Konzentration auf den redebegleitenden *discours citant* erfolgte, weil dieser deutlich häufiger vorkommt als die Redeinleitung und damit viel mehr aussagekräftiges Material für die quantitative Auswertung liefert als diese. Außerdem eignet er sich aufgrund seiner durch die Inversion begründete Formelhaftigkeit hervorragend für Suchanfragen mittels Regulärer Ausdrücke.

Während der *discours cité* bzw. der *discours direct* Gegenstand zahlreicher Überlegungen und Untersuchungen war und immer noch ist, hat der *discours citant* zu keinem Zeitpunkt eine ähnlich große Aufmerksamkeit auf sich gezogen. Obwohl letzterer durch seine metasprachliche Funktion wesentlich zur Charakterisierung der sprechenden Figuren beiträgt, scheint die Beschäftigung mit dem Thema kaum attraktiv zu sein, so dass es nur wenige Einzelbetrachtungen dazu gibt, die obendrein aufgrund ihres Entstehungszeitpunktes vor dem Informationszeitalter bis auf eine Ausnahme rein qualitativer Natur sind. Für Linguisten und Literaturwissenschaftler hält das Internet inzwischen eine kaum überschaubare Fülle von Materialien in elektronischer Form zur Untersuchung bereit. Sobald es jedoch darum geht, breiter angelegte Analysen durchzuführen, stößt man sehr schnell an Grenzen, da in der Regel passende Werkzeuge fehlen, um die Datengrundlage entsprechend zu bearbeiten. Dies trifft auch auf eine quantitative, korpusbasierte Untersuchung des redebegleitenden *discours citant* zu. Zwar gibt es eine Vielzahl elektronischer Romanfassungen, die sich als Datengrundlage anbieten, aber nach elektronischen Hilfsmitteln, d.h. einer auf den Untersuchungsgegenstand genau zugeschnittenen Software, sucht man vergeblich. Bereits in prä-informatischer Zeit wurde darauf hingewiesen, dass eine genauere Analyse der *verba communicandi*

von Interesse sei, sich der Aufwand dafür – das (zum damaligen Zeitpunkt manuelle) „Durchforsten“ unzähliger Romane – aber kaum lohne.¹ Diese Lücke schließt die Verfasserin der vorliegenden Arbeit, indem sie ein eigens für diesen Zweck konzipiertes und entwickeltes Programm, DCAnalyse, zur Verfügung stellt. Nun könnte man argumentieren, dass der Programmieraufwand groß ist im Vergleich zum Umfang der gewonnenen Ergebnisse. Wenn man jedoch berücksichtigt, dass die durchgeführten Analysen ohne Zuhilfenahme des Computers gar nicht erst möglich gewesen wären und das Programm für weitere linguistische und literaturwissenschaftliche Studien genutzt werden kann, ohne dass erneut viel Zeit in die Erschließung des Materials investiert werden muss, ist die Mühe durchaus gerechtfertigt. Eine weitgehend automatisierte Datenverarbeitung durch den Computer hat außerdem den entscheidenden Vorteil, dass auch große Datenmengen ermüdungsfrei bewältigt werden können. Im ungünstigsten Fall wird einfach nur mehr Rechenzeit benötigt. Das Suchen und Finden bestimmter Textstellen erledigt die Maschine konsequent, d.h. so (gut oder schlecht), wie der/die Programmierer(in) dies vorgibt.

DCAnalyse kann mit HTML- oder einfachen Textdateien als Input umgehen. Es erfolgen automatisch mehrere Verarbeitungsschritte, d.h. es werden Formatierungen vorgenommen und Textstellen mit direkter Rede anhand graphischer Marker erkannt, extrahiert und in eine Ausgabedatei exportiert. Nach der Unterscheidung zwischen redееinleitenden, redēbegleitenden und alleinstehenden Sätzen werden die in der Redееinleitung und der Redēbegleitung enthaltenen *verba communicandi* identifiziert und für spätere Auswertungen gespeichert. Außerdem wird detailliert untersucht, welcher Konstruktions-typ innerhalb des *discours citant* vorliegt. Zu allen Analysen wird für jede untersuchte Textdatei umfangreiches Datenmaterial im HTML- und Textformat angefertigt, das die Ergebnisse illustriert und als Grundlage für weitere, hier nicht angestellte Untersuchungen dienen kann.

Betrachtet man die Ergebnisse, die DCAnalyse für die 70 im Arbeitskorpus enthaltenen Texte – der Untersuchungszeitraum beträgt 175 Jahre – produziert, ergibt sich ein differenziertes Bild der diachronen Entwicklung vor allem der Redēbegleitung, das den sich allmählich vollziehenden Wandel der Diskurstradition widerspiegelt. Während sowohl das Set der verwendeten Redeverben als auch die Ausgestaltung der Redēbegleitung und Redееinleitung ganz zu Beginn des Untersuchungszeitraums noch überschaubar bzw. einfach gehalten sind, so wird daraus im Laufe der Zeit ein komplexes System mit einer Vielzahl von möglichen Konstruktionen.

¹ Vgl. Vivienne Mylne (1994), S. 42.

Im Rahmen der Auswertung der mithilfe von DCAnalyse ermittelten Daten wurden vor allem allgemeine Eigenschaften der direkten Rede sowie verschiedene Merkmale des redebegleitenden *discours citant* erörtert.

Bezüglich allgemeiner Eigenschaften der Figurenrede lassen sich folgende Entwicklungen ausmachen: Der Anteil der direkten Rede am Erzähltext liegt für die erste Korpus-hälfte auf einem deutlich höheren Niveau als für die zweite Korpus-hälfte. Dieser Umstand kann als Indiz dafür angesehen werden, dass die Bedeutung der direkten Rede im Text abnimmt und an ihre Stelle zunehmend Neuerungen wie der *discours indirect* und der *discours indirect libre* treten. Unterscheidet man weiter nach redeeinleitenden, redebegleitenden und alleinstehenden Sätzen, sind auch dort interessante Entwicklungen auszumachen. Die Redeeinleitung spielt in der Mehrzahl der Korpus-texte nur eine untergeordnete Rolle, da sie in mehr als der Hälfte der Texte weniger als 10% beträgt. Für die nach 1850 entstandenen Romane ist allerdings eine leichte Zunahme des redeeinleitenden *discours citant* zu beobachten. Die Redebegleitung hingegen dominiert in mehr als jedem zweiten Text mit einem Anteil von mindestens 50%, und in immerhin mehr als zwei Dritteln der Romane liegt ihr Anteil bei mindestens 30%. Ab 1879 ist eine deutliche Abnahme der Redebegleitung vor allem zugunsten alleinstehender Turns, aber auch der Redeeinleitung, erkennbar. Das ist möglicherweise ein Anzeichen dafür, dass sich die Lesegewohnheiten dahingehend verändert haben, dass das Gesagte automatisch den richtigen Personen zugeordnet wird, solange die Zahl der Sprecher überschaubar bleibt, und auf die Redekennzeichnung häufig verzichtet werden kann.

In einem weiteren Schritt wurde der redebegleitende *discours citant* auf bestimmte Merkmale hin untersucht. Zunächst wurde ermittelt, welche *verba communicandi* bevorzugt in der Redebegleitung eingesetzt werden. Dabei zeigte sich, dass nicht nur der Zeitpunkt der Veröffentlichung bei der Wahl der Rede-Verben eine Rolle spielt, sondern auch sehr stark die individuellen Vorlieben des Autors zum Tragen kommen. Besonders gut erkennbar ist dies anhand der sehr produktiven Romanciers Honoré de Balzac und Émile Zola, von denen je 16 Romane ins Korpus aufgenommen und untersucht wurden. Beide verfügen über ein bestimmtes Repertoire an Rede-Verben, die bis auf wenige Ausnahmen in allen Texten Eingang in die Redebegleitung finden.² Während die Redebegleitung zu Beginn des Untersuchungszeitraums mit maximal zwölf Rede-Verben noch kaum ausgestaltet wird und quasi nur die Funktion der Sprecherkennzeichnung übernimmt, wird sie nach 1823 nicht zuletzt aufgrund zusätzlicher, vermehrt auch ex-

² Hierbei handelt es sich um ein Stilmerkmal, das auf die meisten Autoren zutreffen dürfte. Dies für alle im Korpus enthaltenen Autoren nachzuweisen, ist allerdings schwierig, da einige nur wenige Romane publiziert haben und zudem nicht alle Romane in elektronischer Fassung vorliegen.

pressiver *verba communicandi* deutlich variantenreicher gestaltet. Wie viele Redeverben letztlich in einem Text enthalten sind, hängt von mehreren Faktoren ab, nämlich von der Textlänge und vom Anteil von direkter Rede sowie dem darin enthaltenen Anteil von redebegleitenden Sätzen. Von Autor zu Autor ergeben sich immer wieder größere Unterschiede. So sind die beiden untersuchten Romane von Eugène Sue die längsten im Korpus, enthalten einen hohen Anteil von direkter Rede und weisen innerhalb des gesamten Korpus die meisten Redeverben auf³. Bei Jules Verne hingegen verhält es sich ganz anders. Der Roman *L'île mystérieuse* (1875) ist zwar um ein gutes Drittel länger als *Vingt-mille lieues sous les mers* (1869), hat aber bei einem vergleichbaren Anteil von direkter Rede⁴ und einem mehr als doppelt so großen Anteil von redebegleitenden Sätzen⁵ trotzdem vier VC-Types weniger als der kürzere Text. Bei Victor Hugo tritt außerdem ein klarer Bruch hinsichtlich der Verwendung der Redeverben zutage: Während in seinem Spätwerk *Les misérables* (1862), das knapp vier bzw. drei Jahrzehnte später erschien als die beiden frühen Romane *Han d'Islande* (1823) und *Notre-Dame de Paris* (1831), sowohl der Anteil der direkten Rede als auch der Anteil der Redebegleitung deutlich geringer ausfällt als in den beiden älteren Texten, sind trotzdem mehr VC-Types enthalten. Wenig überraschend ist hingegen, dass *dire* in fast allen untersuchten Romanen das häufigste *verbum communicandi* innerhalb der Redebegleitung ist, meist sogar mit sehr deutlichem Abstand zum zweithäufigsten. Die einzige Ausnahme bilden die vier Texte von Jules Verne, in denen *répondre* das größte Vorkommen aufweist, ohne dass jedoch ein zu erwartendes komplementäres Kommunikationsverb wie *demander* oder entsprechende Synonyma in ebenso großer Zahl vorhanden wären.

Es wurde außerdem der Frage nach der Verteilung und Entwicklung der in der Redebegleitung vorkommenden Konstruktionstypen⁶ nachgegangen. Der redebegleitende *discours citant* kann zwar vielfältig gestaltet werden, aber bedingt durch die Inversion von Subjekt und Verb sind die Möglichkeiten auch nicht unbeschränkt. Bestimmte Konstruktionen kehren daher immer wieder. Die Zahlen für die einzelnen Autoren und Romane fallen sehr unterschiedlich aus, und selbst in verschiedenen Werken ein und desselben Autors kommt es mitunter zu großen Schwankungen. Für alle Romane gilt jedoch, dass

³ Für beide Romane wurden ca. 580.000 Tokens gezählt. *Les mystères de Paris* (1842) besteht zu 64,1% aus direkter Rede, die wiederum zu 31,1% redebegleitende Sätze sowie 75 VC-Types enthält. Der Anteil von direkter Rede beträgt in *Le juif errant* (1845) 53,7%. Der redebegleitende *discours citant* beläuft sich auf 48,2%, und es werden 85 VC-Types eingesetzt.

⁴ Anteil der direkten Rede in *L'île mystérieuse* (1875): 23,7% bzw. in *Vingt-mille lieues sous les mers* (1869): 25,7%.

⁵ Anteil der redebegleitenden Sätze in *L'île mystérieuse* (1875): 75% bzw. in *Vingt-mille lieues sous les mers* (1869): 34,4%.

⁶ Erläuterung der Konstruktionstypen siehe Übersicht auf S. 113.

der Basistyp T1, «*dit-il*», der häufigste ist. Beträgt der Anteil von T1 mehr als 50%, was auf mehr als die Hälfte aller Texte zutrifft, ist die Verwendung weiterer Konstruktionstypen eingeschränkt. Liegt er darunter, ist die Redebegleitung in diesen Texten deutlich elaborierter. Über die Zeit hinweg sind außerdem gewisse Verschiebungen auszumachen. Während in der ersten Korpuschälfte neben T1 vor allem die Dativkonstruktion (T3e), die *gérondif*-Konstruktion (T3b) und das Präpositionalsyntagma (T3c) populär sind, geht deren Gebrauch in der zweiten Korpuschälfte deutlich zurück, wovon vor allem die Adverbialkonstruktion (T3a), aber auch der Basistyp selbst profitiert.

Im Anschluss daran wurde ermittelt, welche Redeverben innerhalb der Konstruktionstypen eingesetzt werden. Es zeigte sich, dass die verwendeten VC-Typen in den einzelnen Autorengruppen von Konstruktionstyp zu Konstruktionstyp nur geringfügig variieren. Das gilt auch für alle Korpusstexte zusammengenommen. Nur für die Dativkonstruktion (T3e) ergibt sich, unabhängig von der Autorengruppe, eine andere Zusammensetzung der Verben. *Dire* wird in der Dativkonstruktion noch häufiger als in den übrigen Konstruktionstypen gebraucht, wohl hauptsächlich weil das verwendete *verbum communicandi* die geforderte Valenz aufweisen muss. Viele Verben kommen daher von vorneherein nicht für den Gebrauch in Frage. Zu den zehn häufigsten Redeverben gehören sechs Verben, die für alle Konstruktionstypen nachgewiesen wurden, und zwar *dire*, *demander*, *s'écrier*, *faire*, *répondre* und *reprendre*. Daran ändert sich auch im letzten Drittel des Untersuchungszeitraums nichts, obwohl sich ebenfalls abzeichnet, dass die Vielfalt der *verba communicandi* zunimmt. Bemerkenswert ist auch, dass sich die Zahl der Redeverben, die die Stimmqualität beschreiben, trotz größerer Verbenvielfalt ab der zweiten Korpuschälfte quasi nicht verändert. Auf das gesamte Korpus bezogen sind dies nur die Verben: *crier*, *s'écrier* und *murmurer*.

Um beurteilen zu können, welche Autoren sich um einen besonders variantenreichen Einsatz von *verba communicandi* bemühen, wurde die Type-Token-Relation für die Redebegleitung für jeden im Korpus enthaltenen Roman berechnet. Die TTR der *verba communicandi* ist per Definition umso höher, je näher die Anzahl der VC-Tokens und die Anzahl der VC-Typen beieinanderliegen, und sie ist umgekehrt umso niedriger, je mehr VC-Tokens im Vergleich zu VC-Typen enthalten sind.⁷ Die Anzahl der VC-Tokens kann zwar beliebig zunehmen, die Zahl der möglichen VC-Typen jedoch nicht proportional dazu anwachsen. Zu Beginn des Untersuchungszeitraums, d.h. zwischen 1747 und 1787, beträgt die TTR in den ersten sieben Texten 10% und mehr, was sich darauf zurückführen lässt, dass die direkte Rede selten ist, weshalb sich wenige VC-Tokens auf wenige

⁷ Zur Illustration dieses prinzipiellen Sachverhalts siehe die Abbildungen 4.28 und 4.29.

VC-Typen verteilen. Zwischen 1823 und 1862 liegt die TTR am niedrigsten und übersteigt nur in neun Fällen überhaupt die Fünf-Prozent-Marke. Danach ist bis 1887 keine eindeutige Entwicklung erkennbar, außer dass sich die streuenden Werte insgesamt auf einem höheren Niveau bewegen. Ab 1892 lässt sich schließlich für fast alle Romane eine deutliche Zunahme der TTR beobachten, die bis zum Ende des Untersuchungszeitraums anhält. Dieser Anstieg hat mehrere Ursachen: Einerseits nimmt der Anteil der direkten Rede im Vergleich zum Erzähltext ab, während die direkte Rede gleichzeitig seltener mit einem redebegleitenden *discours citant* in Verbindung steht. Ein geringerer Anteil von direkter Rede, verbunden mit einem Rückgang der Redebegleitung, hat weniger Dialogstellen bzw. eine geringere Anzahl von VC-Tokens zur Folge. Da die Anzahl der VC-Typen jedoch nicht in gleichem Maße dazu abnimmt, sondern weitgehend konstant bleibt, steigt die TTR. Dies bestätigt die bereits zuvor bei der Betrachtung der *verba communicandi* gewonnene Erkenntnis, dass deren Einsatz zum Ende des Untersuchungszeitraums hin variantenreicher wird, einzelne Verben, allen voran *dire*, jedoch nach wie vor zu den häufigsten gehören.⁸ Geht man noch einen Schritt weiter und betrachtet die TTR der *verba communicandi* für die Konstruktionstypen, so lässt sich feststellen, dass sie sich für die sehr häufigen Konstruktionstypen (Adverbialkonstruktion (T3a), *gérondif*-Konstruktion (T3b) und Dativkonstruktion (T3e)) vergleichbar zu der TTR für die gesamte Redebegleitung entwickelt und nur die selteneren Konstruktionstypen (Konstruktion mit Adjektiv oder *participe passé* (T2a) bzw. das Präpositionalsyntaxagma) keinem eindeutigen Trend folgen.

Überprüft wurde auch die von Vivienne Mylne (1994) aufgestellte These, dass sich vor allem die Verben aus der *gérondif*-Konstruktion, die ein Redeverb semantisch erweitern, selbst zu *verba communicandi* entwickeln. Dazu wurde für jeden Roman die Häufigkeit sowohl der innerhalb der *gérondif*-Konstruktion verwendeten Verben als auch der entsprechenden konjugierten Verbformen ermittelt und auf der Zeitachse einander gegenübergestellt (siehe Kapitel 4.2.5). Daraus ergibt sich einerseits, dass die Anzahl der expressiven Redeverben zwar mit fortschreitender Zeit zunimmt, andererseits aber die konjugierte Form des jeweiligen Verbs in der Regel ab dem Zeitpunkt ihres erstmaligen Auftretens überwiegt. Die *gérondif*-Konstruktion ist zudem im Vergleich zu den konjugierten Formen viel zu selten, als dass sie als Auslöser für eine so weitreichende Veränderung gelten könnte. Vivienne Mylnes These ist also so nicht haltbar. Es ist zwar richtig,

⁸ Zu diesem Schluss kommt auch Danielle Coltier (1989): «*Cependant, il convient de noter [...] que la variété de ces verbes n'est pas toujours aussi grande qu'on pourrait le penser a priori; si les romans réalistes / naturalistes en emploient plus que «les romans de l'âge classique» [...], l'abondance n'est pas signe de variété effective*» (S. 88).

dass die expressiven Verben ab der zweiten Korpus­hälfte zunehmen und gleichzeitig die Frequenz von *dire* als *Passepartout*-Redeverb etwas abnimmt, aber daraus kann nicht gefolgert werden, dass dies über den Umweg über die *gérondif*-Konstruktion geschieht. Vielmehr dürfte dies auf neue literarische Konventionen sowie idiolektale Vorlieben der Autoren zurückzuführen sein. Die Zusammensetzung der Redeverben verändert sich außerdem keineswegs dahingehend, dass die Bedeutung von *dire* zu irgendeinem Zeitpunkt ernsthaft in Frage gestellt würde.

Der abschließende Vergleich mit vorangegangenen Arbeiten verdeutlichte nochmals die Vorzüge des für diese Arbeit eingesetzten Verfahrens gegenüber der manuellen Datengewinnung zu prä-informatischer Zeit. Zum einen lässt sich die Plausibilität von Ergebnissen nur beurteilen, wenn die Methode transparent ist, zum anderen sollte man von Hand erzeugte Ergebnisse stets mit einem gesunden Misstrauen begegnen, da trotz größter Sorgfalt Fehler entstehen, die ab einer bestimmten Datenmenge nicht mehr nachvollziehbar sind. Der Computer bzw. Reguläre Ausdrücke sind bei ihrer Suche „gierig“ und fördern ohne Ausnahme alles zutage, was ihm bzw. ihnen aufgetragen wurde.

Die vorliegende Arbeit hat sich mit grundlegenden Aspekten des redebegleitenden *discours citant* auseinandergesetzt. Denkbar sind viele weitere, vor allem auch qualitativ ausgerichtete Fragestellungen, die den Rahmen dieser Arbeit deutlich gesprengt hätten und somit von anderen geleistet werden müssen. So wäre von linguistischer Seite äußerst interessant zu ermitteln, wie das zu jedem realen Gespräch gehörende Körperverhalten genau im *discours citant* abgebildet wird. Dies wurde im Zuge dieser Arbeit ansatzweise unternommen, indem versucht wurde, die das Redeverb erweiternde *gérondif*-Konstruktion automatisch zu kategorisieren.⁹ Es zeigte sich jedoch, dass dies ohne erhebliche weitere Programmierarbeit bzw. ohne Zuhilfenahme ergänzender Software nicht zu leisten ist.¹⁰ Der Aufwand erscheint durchaus lohnenswert, da die Ergebnisse – ähnlich wie in dieser Arbeit – sehr interessant und vielleicht auch überraschend sein dürften.

Wie eingangs bereits erwähnt wurde, kann DCAnalyse auch zur Bearbeitung literaturwissenschaftlicher Themen genutzt werden. Das Programm kann ohne weitere Vorarbeit Material zur Erörterung diverser Fragen bereitstellen. So ließe sich beispielsweise –

⁹ Durch eine Unterscheidung nach den Kriterien Gestik, Mimik, Stimme, Augen, Bewegung, Sprecherwechsel und Art der Wahrnehmung wurde eine Kategorisierung versucht. Die (provisorischen) Ergebnisse finden sich nach dem Einsatz von DCAnalyse in der Datei `./output/dateiname/gerondif_konstruktionen__dateiname.[txt|html]`.

¹⁰ Anders als die Redeverben, die in ihrer Zahl aufgrund ihres Einsatzgebietes begrenzt sind, unterliegen die innerhalb der *gérondif*-Konstruktion verwendeten Verben so gut wie keiner Beschränkung. Die Schwierigkeit liegt also in der Zusammenstellung einer kategorisierten Liste, die alle innerhalb der *gérondif*-Konstruktion möglichen Verben erfasst.

auch in größerem Umfang und nicht auf einzelne Werke beschränkt – klären, welche Figuren innerhalb des *discours citant* welche Bezeichnungen erhalten, wie häufig welche Personen zu Wort kommen und welche Personen bevorzugt welche Verben benutzen.¹¹ Dieses Datenmaterial ebenfalls automatisiert auf die entsprechende Fragestellung hin zu erschließen und damit das Programm um eine weitere Komponente zu erweitern, steht ebenfalls jedem interessierten Benutzer offen.

¹¹ Diese Aspekte wurden anhand einzelner Romane bereits erörtert. Siehe dazu die Beiträge von Prince (1980a), Prince (1980b) und Halina Suwala (1987).

Anhang A

Verwendete *verba communicandi* in der Redebegleitung

A.1 Ermittelte VC-Types und Anzahl der VC-Tokens

A.1.1 Übersicht VC-Types: Bernardin de Saint-Pierre und Voltaire

	SPI_PAUVIR	VOL_BLANCN	VOL_CANDID	VOL_INGENU	VOL_MEMNON	VOL_PRINCB	VOL_ZADIG
Tokens ges.	40432	14349	34259	25012	21152	26588	27441
VC-Tokens ges.	45	85	282	69	22	67	135
VC-Types (rb)	6	8	7	9	6	7	12
VC-Tokens (rb)	27	79	270	61	21	60	124
ajouter							2
commencer			1				
continuer		2	1			1	1
convenir	1						
crier				1			1
demander	1					2	2
dire	14	53	214	36	13	38	73
faire		1		1	1		
insister							1
juger							1
offrir							1
poursuivre				1			
promettre		1					
reconnaître							1
refuser				1			
reprendre	6	3	6	2	2	3	10
répliquer		2	7	7	2	1	
répondre	3	12	30	6	2	11	25
écrier	2	5	11	6	1	4	6

Tabelle A.1: Anzahl der VC-Types und -Tokens in der Redebegleitung bei **Voltaire** und **Bernardin de Saint-Pierre**. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

A.1.2 Übersicht VC-Types: Victor Hugo, George Sand, Stendhal und Eugène Sue

	HUG_HANISLA	HUG_MISERA	HUG_NODAME	SND_CONSUE	SND_FCHAMP	SND_INDIAN	SND_MAREDI	STE_CHARTR	STE_ROUGEN	SUE_JUIFER	SUE_MYSTER
Tokens ges.	126947	557192	181695	360571	48110	98587	39781	194913	185712	578457	581180
VC-Tokens ges.	1206	2142	1012	2027	250	436	180	855	965	5193	3855
VC-Types (rb)	33	47	33	54	14	19	10	25	17	85	75
VC-Tokens (rb)	1136	1526	970	1909	222	394	176	689	893	4501	3631
accepter				1						2	
accorder											1
accuser				1							
achever	1	1								2	1
admettre										1	
affirmer	1										
ajouter	36	27	23	57		9	2	46	63	281	191
amuser		2								1	2
annoncer										3	1
appeler		1	1	2				1	2	4	5
apprendre						1				2	2
appuyer										1	
articuler				1						1	1
assurer	3		1	1						1	1
attester	1										
avancer										1	
aviser				1							
avouer								1		1	1
balbutier	4	2	5								2
bégayer		1	2								
calculer										1	
calomnier											1
causer										1	2
chanter				5							2
chantonner										1	1
charger				1						4	3
chasser				1							
chuchoter			1								
commander		1								1	
commencer		1		1					1	2	2
compter										1	1
conclure				1							
condamner	1			1						1	1
conseiller										1	1
consentir										1	
conter		1									
continuer	29	7	10	4		4		7	26	9	6
convenir										1	
corriger				1							2
couper	1										
crier	35	76	60	18	1	2		26	6	55	54
demander	60	80	38	36	2	3	1	5		101	129
dessiner	1										
dicter											1
dire	448	718	418	827	132	223	103	371	535	2205	1770
distinguer										1	
décider		1		1						2	2
déclarer	1	1		1				1		1	1
défendre				1							
...

A.1 Ermittelte VC-Types und Anzahl der VC-Tokens

Tabelle A.2 – Fortsetzung

	HUG_HANISLA	HUG_MISERA	HUG_NODAME	SND_CONSUE	SND_FCHAMP	SND_INDIAN	SND_MAREDI	STE_CHARTR	STE_ROUGEN	SUE_JUIFER	SUE_MYSTER
dénoncer										2	2
dévoiler											1
emporter											1
empresser	1										
engager										1	
enseigner				3							
entendre				2						1	1
examiner						1					2
exclamer		1			2						
expliquer		1								3	
exposer	1									1	1
exprimer				1							1
extasier										1	
faire	3	80	4	7	22	1	2	6	5	32	19
fredonner		1									
glapir		1	1								
glisser		1									
grincer											1
grogner			1								
grommeler	1	13	13	1							
gronder				2							
hasarder										2	
hurler		3	3			1		1			
informer										1	
inquiéter				1							
insister								1			
instruire										3	1
interroger											1
interrompre	40	10	7	7		4				1	
jeter	1	2	2	4			1	1		4	4
juger				2				1		1	
jurer		2		4							
lire								1			1
louer								1			
lâcher											1
menacer				1						1	
montrer		2		1	1				2	2	3
murmurer	30	24	10	8		1				25	37
noter										1	
négociier											1
objecter			1	1							
observer	1	13	14	8	1		1				
offrir	1							1		3	1
opposer		1									
ordonner											2
parler		2	1	2				3	1	6	6
peindre										1	
penser	3	22	10	27	2	18	1	5	53	15	16
permettre		1		1						1	
persuader										1	
plaindre										1	
pleurnicher											1
poursuivre	36	10	22	2		2		2		2	1
presser										2	
prier	2									1	1
prononcer				1						2	1
proposer										1	1
proscrire		1								1	
prouver										1	3
présenter								1		2	1
...

Tabelle A.2 – Fortsetzung

	HUG_HANISLA	HUG_MISERA	HUG_NODAME	SND_CONSUE	SND_FCHAMP	SND_INDIAN	SND_MAREDI	STE_CHARTR	STE_ROUGEN	SUE_JUIFER	SUE_MYSTER
prétendre										1	
prévenir										1	1
psalmodier											1
raconter										1	4
railler			1							1	2
rappeler			1	2		1					1
recommander											1
reconnaître		1									1
redire										1	
regretter										1	
relever	2	1								2	1
remarquer			1	1					1		1
repartir										1	1
repousser										1	1
reprendre	121	165	134	216	12	24	17	35	48	693	455
revenir	2	1			1					4	5
riposter		3			1						4
réjouir										1	
répliquer	21	11	11	21	7			15	7		2
répondre	99	127	113	428	36	67	40	36	56	313	195
répéter	50	15	10	4		3		12	13	42	48
résumer				1							
révéler				1							
sangloter		2									
siffler				1							
signifier		1									
songer										1	1
souffler					2						
soupirer		1	2								
supplier										1	
trionpher				1						1	
vanter						1					
vociférer			1								
éclairer										2	
éclater										1	
écrier	99	87	48	183		28	8	109	72	617	607
écrire									2	2	
égosiller				1							
énoncer										1	

Tabelle A.2: Anzahl der VC-Types und -Tokens in der Redebegleitung bei **Victor Hugo, Stendhal, George Sand und Eugène Sue**. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

A.1.3 Übersicht VC-Types: Honoré de Balzac

	BAL_CESBIR	BAL_COLCHA	BAL_COUSBE	BAL_COUSPO	BAL_CUREVI	BAL_DUCHES	BAL_EUGENI	BAL_FEMMET	BAL_FILLEY	BAL_LYSVAL	BAL_MEDCAM	BAL_PAYSAN	BAL_PBOURG	BAL_PEAUCH	BAL_PGORIO	BAL_TENAFF
Tokens ges.	105790	24346	155136	111169	93385	53595	67164	69908	28451	106941	89493	116835	64824	94940	96345	75424
VC-Tokens ges.	933	271	1671	972	561	280	684	502	178	398	539	1044	576	649	1065	718
VC-Types (rb)	20	13	23	19	23	13	20	13	12	17	37	30	15	37	26	20
VC-Tokens (rb)	906	250	1583	919	530	250	611	448	160	288	475	984	525	597	1021	671
accorder											1				1	
accoucher											1	1				1
accuser											1					
achever	1													1		
adresser															1	
ajouter	7	3	39	21	6	9	12	18	2	7	13	15	13	21	13	10
amuser														1		
annoncer							2									
appeler											1			1		1
apprendre											1					
appuyer								1						1	1	
avancer					1											
avertir											1					
beugler												1				
calculer	1															
causer				1		1						1		1		1
chanter							1							1	2	
chantonner															1	
commander											1					
commencer				1					1			2				
communiquer							1									
compléter														1		
compter														1		
condamner															1	
confirmer					1											
congédier											1					
consentir				1												
consoler										1						
continuer			1		1									1		
convoquer											1					
crier	14	2	21	9	3	2	26	10	4	9	20	20	8	15	42	6
demander	17	10	101	57	28	12	24	31	4	9	21	50	21	15	27	41
dessiner				1											1	
dire	708	137	841	523	298	142	377	182	103	155	236	530	283	218	664	403
discuter											1					
décider															1	
déclarer														1		
dénoncer																1
désespérer	1													1		
développer			1													
empresser												1				
engager		1			1											
entendre											1	1				
estimer												1				
examiner														2		1
exiger								1								
expliquer					1			1	1				1			
exposer												1				
exprimer											1					
faire	8	1	12	7	6	4	4			7	13	5	9	8	20	5
formuler														1		
...

Anhang A: Verwendete verba communicandi in der Redebegleitung

Tabelle A.3 – Fortsetzung

	BAL_CESBIR	BAL_COLCHA	BAL_COUSBE	BAL_COUSPO	BAL_CUREVI	BAL_DUCHES	BAL_EUGENI	BAL_FEMMET	BAL_FILLEY	BAL_LYSVAL	BAL_MEDCAM	BAL_PAYSAN	BAL_PBOURG	BAL_PEAUCH	BAL_PGORIO	BAL_TENAFF
glisser													1		1	
grogner																
implorer	1															
indiquer					1						1					
insister											1					
instruire														1		
interdire			1													
jeter	1				1					2						
juger	1															
jurer			2				1			1						
lancer												1				
lire		1									1					
menacer																
montrer							1				1	2		1		
moquer							1									
observer																
offrir	1				1						1	1		1		1
opposer											2					
parler	1		1	1			1		2		1	1		4	1	
peindre																
penser	11	1	7	5	3	2	10	2	1	3	4			3	4	6
permettre			1													
poursuivre												1				
prescrire											1					
prier																
pronostiquer										1					1	
proposer											1		1	1		
prouver			1		2											
présenter														2		
prétendre			1													
prévenir				1												
raconter			1								1					
radoter											1					
raisonner	1															
rappeler					1							1			1	
rapporter				1						1						
recommander					2											
reconnaître											1					
refuser											2					
regretter							1									
relever			1											1		
remarquer												1				
reprendre	45	35	167	88	43	27	37	52	15	42	53	92	34	84	89	48
reprocher						1										
représenter											1	1	1			
revenir							1				1					
réciter																
réitérer										1						
répandre	1															2
répliquer		3	47	15	8	3	5	9			2	29	7	14	4	8
répondre	34	27	208	108	68	29	77	72	18	34	62	141	69	73	72	69
répéter	4	2	12	7	2		2	8	3	1	1	5	1	7	9	2
révéler														1		
signifier												1				
songer												1	1	1		
souffler															1	
traduire												1				
traiter														1		
écrier	48	27	115	71	51	17	27	61	6	13	22	75	75	108	61	63
écrire			1	1										1	1	1
...

Tabelle A.3 – Fortsetzung

	BAL_CESBIR	BAL_COLCHA	BAL_COUSBE	BAL_COUSPO	BAL_CUREVI	BAL_DUCHES	BAL_EUGENI	BAL_FEMMET	BAL_FILLEY	BAL_LYSVAL	BAL_MEDCAM	BAL_PAYSAN	BAL_PBOURG	BAL_PEAUCH	BAL_PGORIO	BAL_TENAFF
émerveiller			1													

Tabelle A.3: Anzahl der VC-Types und -Tokens in der Redebegleitung bei **Honoré de Balzac**. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

A.1.4 Übersicht VC-Types: Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert

	DUM_CHRIST	DUM_COLLIE	DUM_MOUSQU	DUM_TULIPE	FLA_BOUVAR	FLA_BOVARY	FLA_EDUCAT	FLA_SALAMM	GON_GERMIN	VER_ILEMYS	VER_LIEUES	VER_TERLUN	VER_TOURMO
Tokens ges.	505014	265042	239109	76269	93354	120479	150090	105524	69443	210369	149385	55824	72060
VC-Tokens ges.	5882	3047	2914	665	222	732	699	210	165	1773	593	392	553
VC-Types (rb)	67	63	41	31	24	36	38	23	16	22	26	16	15
VC-Tokens (rb)	5821	3017	2866	652	184	622	532	149	126	1708	568	371	521
accepter		1											
accorder			1										
accuser					1								
achever	1	2		1									
admirer	1	1											
adresser	1												
affirmer		1											
ajouter	74	54	28	8	2	15	8	1		53	19	1	22
amuser							2		1				
annoncer	1	1		1	1						1		
appeler	7	1				1	1						
apprendre											1		
apprécier		1											
appuyer	2	1											
articuler	1	3				1	1						
assurer	1	1											
attaquer						1					1		
avancer		1											
avouer	2		1		1		2						
balbutier	24	15	7	4		3	5		1				2
calculer	1												
cancaner							1						
causer	1									1			
chanter									1				
charger	2		2							1			
chuchoter								1					
commander							1						
commencer	4	1				1		1	1				
communiquer		1											
conclure											1		
concéder						1							
condamner													1
confier	1												
conférer	1												
consoler		1				1		1					
constater												1	
...	1	...

Anhang A: *Verwendete verba communicandi in der Redebegeleitung*

Tabelle A.4 – Fortsetzung

	DUM_CHRIST	DUM_COLLIE	DUM_MOUSQU	DUM_TULIPE	FLA_BOUVAR	FLA_BOVARY	FLA_EDUCAT	FLA_SALAMM	GON_GERMIN	VER_ILEMYS	VER_LIEUES	VER_TERLUN	VER_TOURMO
conter	2		2										
continuer	199	107	112	34		12	2	1					
convaincre			1										
couper		1						1			1		1
crier	33	42	29	26	2	14	5	6	3	9	3	2	2
demander	534	114	186	61	1	29	11	6		173	77	29	52
dessiner	1												
dire	3199	1414	1578	271	93	249	278	69	64	422	99	80	116
décider			1										
déclarer							1	1					
délirer		1											
dénigrer					1								
emporter	1	1											
engager					1								
entendre	1			1									
exalter							1						
examiner											1		
exclamer	1	3		1		9	2	1					
excuser							1						
exiger											1		
expliquer	2												
exprimer	1												
faire	195	224	58	46	2	52	15	11	30	16	32	9	8
glapir							2						
glisser		2		1									
grincer		1						1					
grogner		1	1										
grommeler		5	1	1		1	1						
gronder		2		1									
gémir				2						1			
hasarder	3	1	1						1				
hurler	1	5	1	8			2			1			
implorer		1					1						
indiquer	1												
inquiéter		1											
insister	2	3		1									
insulter		1											
interroger				3									
interrompre	18	32	19			8			1				
inviter									1				
jeter	6					1							
juger						1			1				
jurer			1										
lancer												1	
lire	1		1		1								
lâcher	1					1							
marmotter			1										
maugréer				2									
montrer			2		2						1		
murmurer	136	106	58	26	1	8	3	2	1	20	7	1	5
objecter	1	7	1		4	7	4						
observer			4			1							
offrir	2	2	1								2		
ordonner	3	2						1					
parler	1	1	4		1	2					2		
penser	9	38	13	1		14	11			6	3		6
permettre	1		2								1		
poursuivre	8	26	1	2		1	1						
presser	1												
...

Tabelle A.4 – Fortsetzung

	DUM_CHRIST	DUM_COLLIE	DUM_MOUSQU	DUM_TULIPE	FLA_BOUVAR	FLA_BOVARY	FLA_EDUCAT	FLA_SALAMM	GON_GERMIN	VER_ILEMYS	VER_LIEUES	VER_TERLUN	VER_TOURMO
prier	2		1										
proférer					1								
promettre	2	1											
proposer			1							1			
prouver	1		1		1								
présenter	4	1								1			
prétendre	1						1						
questionner		1											
raisonner							1						
rappeler	1												
rapporter				1							1		
recommander										1			
reconnaître	1												
redire	1												
regretter		1											
relever										1			
remarquer							2					1	
repousser	1	1											
reprandre	363	70	195	9	24	72	72	7	12	55	37	48	34
revenir	2		2							1			
riposter		3			1					7	16	5	
rugir		2		2									
réitérer		1											
répliquer	42	177	12	6	8	19	27	2		28	33	29	3
répondre	495	164	194	47	8	45	18	14	4	675	180	107	203
répéter	63	53	17	11	2	7	2	3	3	20	7	4	9
résoudre								1			1		
résumer		1											
révéler												1	
siffler								1					
songer		1				3	4	1					
souffler							1						
soupirer	2	7	1	1		5	2						
supplier		1		1									
supposer	1												
tonner								1					
traduire	1												
vanter						1	1						
vociférer	1	1	1				1						
éclater					1								
écrier	346	299	323	72	24	33	38	16	1	215	40	52	57
écrire	1	2											
émettre							1						
étonner	1					1							

Tabelle A.4: Anzahl der VC-Types und -Tokens in der Redebegeleitung bei **Alexandre Dumas, Gustave Flaubert, Edmond und Jules de Goncourt und Jules Verne**. Die Roman-kürzel sind nach Tabelle 3.1 aufzulösen.

A.1.5 Übersicht VC-Types: Émile Zola

	ZOL_ARGENT	ZOL_ASSOMM	ZOL_BETEHU	ZOL_BONHEU	ZOL_CONQUE	ZOL_CUREE	ZOL_DEBACL	ZOL_DRPASC	ZOL_EXCELL	ZOL_GERMIN	ZOL_NANA	ZOL_PAGEAM	ZOL_TERRE	ZOL_THRAQU	ZOL_TRVIL	ZOL_VENTRE
Tokens ges.	155024	169753	136823	158853	123600	113110	202232	122575	136916	177544	152160	109902	177461	70946	187612	118428
VC-Tokens ges.	273	756	280	982	1016	418	485	245	929	655	1018	756	644	214	400	476
VC-Types (rb)	30	33	30	36	42	22	38	25	54	46	36	24	43	21	33	30
VC-Tokens (rb)	246	587	208	784	856	307	369	188	631	514	830	600	509	180	310	370
achever					3	1				1						1
adresser									1							
affirmer	1			1	3		1		2		1			1		2
ajouter		2	3	8	10	1	4	1	7	5	7		8	1	2	7
amuser				2					1							
annoncer			1	2								2				
appeler		2		2	3		2		2		2	1	4	2		
apprendre																1
approuver		1						1								
appuyer		1							2	1			1	2	1	
articuler							1									
assurer				2	2											
aviser													1			
avouer			1		1	2				1	1			1		
balbutier	3	3	1	11	11	4		3	2	6	8	15	2	5	1	1
bégayer	1	13	4	2	8	1	6	3		7	13	7	4		3	
causer									1							
chanter												1			1	1
chuchoter				1	1		1									
clamer											1					
commander										2			1			
commencer		1		4			1	1	6	1	3	1			1	1
compter					1											
conclure	6	1	6	7	3	3	6	4	3	1	3		12	1	6	1
confesser		1											1		1	
confirmer							2									
conseiller															1	
consentir									2							
constater										1						
conter																
continuer	12	15	12	16	27	19	9	10	26	17	12	13	17	6	15	9
couper													1			
crier	19	81	8	55	46	7	37	19	41	66	87	34	60	17	16	13
demander	12	43	21	70	70	24	27	12	41	41	74	77	48	6	25	36
dire	44	181	43	222	278	119	82	47	220	127	283	191	112	58	60	135
décider										1						1
déclamer							1					1				
déclarer	14	9	4	19	6	1	10	2	11	8	18	8	30	1	5	2
défendre									1							
emporter		1		1	1								1			
engager									1							
entendre											1					
entêter									1							
examiner	1															
exiger	1															
expliquer	4	8	2	5	2		5	1	4	2	4		1		5	2
faire	2		1	1	3	1		3	4	3		1	1			
geindre	1															
glisser					1											
grogner		5	3				6			3			3			2
grommeler					1											
...

A.1 Ermittelte VC-Types und Anzahl der VC-Tokens

Tabelle A.5 – Fortsetzung

	ZOL_ARGENT	ZOL_ASSOMM	ZOL_BETEHU	ZOL_BONHEU	ZOL_CONQUE	ZOL_CUREE	ZOL_DEBACL	ZOL_DRPASC	ZOL_EXCELL	ZOL_GERMIN	ZOL_NANA	ZOL_PAGEAM	ZOL_TERRE	ZOL_THRAQU	ZOL_TRVIL	ZOL_VENTRE
gronder			4	1	2		6	1	1	4			4			1
gueuler		5								1	2		8			
gémir	1						1					1	2			
hasarder				1	4					1	1		1			1
hurler	1	5			2		1			10	1		3	1		
indigner											1					
indiquer									1							
inquiéter									1							
insinuer					2											
insister	1				2		1								1	
interroger		1						1								
interrompre	5	5	1	8	9	4	2	3	6	2	5	3	5		3	6
intervenir				1									1			
jeter		1							2	1	1		1		1	
jurer		1	1	1			2	1	1	1	1		2			
lancer					2				1	1			2			
lâcher	1	1	1	1					1	1	2	1	2			
menacer															1	
mentir										1						
montrer									3	2	1				1	
moquer									1							
mugir	1															
murmurer	23	68	19	81	98	34	30	13	63	50	100	76	33	8	32	35
nier									1							
négocié	1															
objecter	1	1										1	1			
offrir													2		2	1
ordonner									1							
parler					1				5		1	2				
penser	1		3		11	4			1	1			1	4	1	3
permettre							2									
plaider							1									
plaindre					1											
plaisanter									1							
poursuivre					1	2			2	2			3			1
prier								1								
proclamer									1							
promettre									1							
proposer							1			1			2		1	
questionner	1															
raconter		2	2	1	2		2		2	1	3	1			2	1
rappeler			2					1								
rapporter										1					1	
recommander					2		1			1					1	
reconnaître									1							
refuser			2						1							
relever							1			1					2	
remarquer					1											1
reprandre	39	40	28	93	64	12	38	26	46	44	61	46	41	20	41	34
revenir					1				2							
ricaner		1			1					1				1		
rugir							1									
réclamer									1				1			
récrier							2			1						
répandre								1								
répliquer			1	1	5	2			1	2	1					
répondre	19	43	10	90	87	44	15	7	64	43	68	50	36	28	27	38
répéter	20	31	12	60	35	7	48	12	24	30	37	48	36	5	32	14
sangloter						1	1			1				1		
songer			1				1									
...

Tabelle A.5 – Fortsetzung

	ZOL_ARGENT	ZOL_ASSOMM	ZOL_BETEHU	ZOL_BONHEU	ZOL_CONQUE	ZOL_CUREE	ZOL_DEBACL	ZOL_DRPASC	ZOL_EXCELL	ZOL_GERMIN	ZOL_NANA	ZOL_PAGEAM	ZOL_TERRE	ZOL_THRAQU	ZOL_TRVIL	ZOL_VENTRE
souffler	1	1	1	1					1		5					
soupirer				1					2				1			1
traiter																
éclairer									1		1					
écrier	9	13	10	9	44	14	11	14	11	15	18	19	11	11	17	18
écrire				1					1							

Tabelle A.5: Anzahl der VC-Types und -Tokens in der Redebegleitung bei **Émile Zola**. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

A.1.6 Übersicht VC-Types: Henri Alain-Fournier, Marcel Prévost, Marcel Proust und Colette

	AFOU_GD_ME	COLE_CHERI	PREV_AUTOM	PREV_VIERG	PROU_GUERM	PROU_OMBRE	PROU_SWANN
Tokens ges.	68607	41235	84290	73461	247669	223725	181851
VC-Tokens ges.	176	335	375	463	1015	623	638
VC-Types (rb)	15	67	27	33	41	28	25
VC-Tokens (rb)	128	227	284	386	584	218	273
achever		2			2		
acquiescer		2					
admirer		1					
adresser						1	
affirmer					1		
ajouter	4		5	14	76	27	29
appeler		2					
apprendre		1					
approuver		1			2		
apprécier							1
assurer						1	1
avancer	1						
avouer		1		1			
balbutier			5	3			
bredouiller		1					
bêler		1					
certifier							1
chanter				1			
chevroter		1					
chuchoter		1	1				
clamer		1					
commander		1		1	1		
commencer					1	1	1
compléter						1	
conclure			1	3	9	2	1
concéder		1					
confier		1					1
conseiller		1				1	
constater		4				1	
conter		1					
continuer	3	3	3	2	9	2	1
...

A.1 *Ermittelte VC-Types und Anzahl der VC-Tokens*

Tabelle A.6 – Fortsetzung

	AFOU_GD_ME	COLE_CHERI	PREV_AUTOM	PREV_VIERG	PROU_GUERM	PROU_OMBRE	PROU_SWANN
crier	1	4			4	1	3
demander	9	14	19	42	44	20	12
dire	68	74	71	76	242	86	141
décider		1					
déclarer			2	5	2	2	1
dénoncer							1
désoler					1		
empresser							1
enchérir					2		
entendre					1		
esclaffer		1					
exclamer				1	1		
excuser		1					
expliquer	1				2		
faire	7	2	59	75	2	5	1
glapir		2					
glisser					1		
grogner							1
grommeler		2			3	1	
gronder		1		1			
haleter		1					
hasarder					1		
hennir		1					
hurler					1		
implorer				1	1		
informer					1		
insister		5	2		2		
interroger		2					
interrompre		5	4	13	13	3	3
intervenir		1					
jeter		3			2	1	
juger		1					
lire					1		
montrer		1					1
murmurer		3	19	19	5	1	
noter		1					
objecter		1	1	1			
observer				8			
opiner			1				
ordonner					2		
parler	1		1				
penser	2	4	20	7	4		
piauler		1					
poursuivre	4	2	5	7	2	1	
prier		1	1	1			
prononcer			1	1			
protester		5			1		1
questionner			4	3	1		
quémander		1					
raconter		1			1		
rappeler		1					
recommander							1
reconnaître		1					
rectifier		1		2			
redire		1					
remarquer		1				1	
reprendre		4	11	16	32	5	9
reprocher		2					
revenir				1			
ricaner				1			
...

Tabelle A.6 – Fortsetzung

	AFOU_GD_ME	COLE_CHERI	PREV_AUTOM	PREV_VIERG	PROU_GUERM	PROU_OMBRE	PROU_SWANN
riposter	1	1	20	52	1	1	
répliquer	24	12	14	14	2	1	2
répondre	1	7	6	2	65	34	35
répéter						1	
sommer					1	1	
songer		10					
souffler		2	1				
souhaiter				1		1	
souligner		1					
soupirer	1	6					5
suggérer		1					
supplier			2	2			
télégraphier		1					
éclater		1					
écrier		4	5	9	39	15	19
écrire		1					

Tabelle A.6: Anzahl der VC-Types und -Tokens in der Redebegleitung bei **Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette**. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

Anhang B

Type-Token-Relation der *verba communicandi*

	VC-Tokens ges. (RB)	VC-Types ges. (RB)	TTR RB	T1: VC-Tokens	T1: VC-Types	TTR T1	T2a: VC-Tokens	T2a: VC-Types	TTR T2a	T3a: VC-Tokens	T3a: VC-Types	TTR T3a	T3b: VC-Tokens	T3b: VC-Types	TTR T3b	T3c: VC-Tokens	T3c: VC-Types	TTR T3c	T3e: VC-Tokens	T3e: VC-Types	TTR T3e	
VOL_ZADIG	124	12	9,7	92	9	9,8	1	1	100	2	2	100	3	3	100	0	0	0	20	2	10	
VOL_MEMNON	21	6	28,6	12	4	33,3	2	1	50	0	0	0	1	1	100	0	0	0	5	4	80	
VOL_CANDID	270	7	2,6	219	6	2,7	3	2	66,7	5	2	40	4	2	50	0	0	0	34	3	8,8	
VOL_BLANCN	79	8	10,1	66	7	10,6	1	1	100	1	1	100	0	0	0	0	0	0	6	2	33,3	
VOL_INGENU	61	9	14,8	44	6	13,6	0	0	0	0	0	0	1	1	100	1	1	100	13	3	23,1	
VOL_PRINCB	60	7	11,7	39	6	15,4	0	0	0	0	0	0	2	2	100	0	0	0	16	3	18,8	
SPI_PAUVIR	27	6	22,2	19	4	21,1	0	0	0	1	1	100	0	0	0	0	0	0	6	1	16,7	
HUG_HANISLA	1136	33	2,9	641	16	2,5	52	15	28,8	76	14	18,4	94	14	14,9	70	14	20	30	8	26,7	
STE_ROUGEN	893	17	1,9	367	13	3,5	30	7	23,3	58	9	15,5	77	8	10,4	50	6	12	211	8	3,8	
BAL_PEAUCH	597	37	6,2	268	17	6,3	9	6	66,7	24	6	25	148	12	8,1	18	6	33,3	56	9	16,1	
HUG_NODAME	970	33	3,4	582	24	4,1	34	11	32,4	78	14	17,9	87	15	17,2	54	10	18,5	32	8	25	
SAN_INDIAN	394	19	4,8	158	11	7	8	3	37,5	20	5	25	70	9	12,9	32	5	15,6	77	2	2,6	
BAL_MEDCAM	475	37	7,8	240	20	8,3	1	1	100	20	8	40	86	7	8,1	18	6	33,3	57	7	12,3	
BAL_DUCHES	250	13	5,2	89	10	11,2	3	2	66,7	9	5	55,6	78	8	10,3	20	5	25	29	5	17,2	
BAL_FILLEY	160	12	7,5	72	9	12,5	4	3	75	4	3	75	34	7	20,6	5	2	40	23	4	17,4	
BAL_CESBIR	906	20	2,2	406	11	2,7	37	5	13,5	42	8	19	174	10	5,7	22	3	13,6	112	3	2,7	
STE_CHARTR	689	25	3,6	241	14	5,8	31	6	19,4	26	9	34,6	63	9	14,3	45	7	15,6	204	11	5,4	
BAL_FEMMET	448	13	2,9	159	11	6,9	4	3	75	32	6	18,8	97	9	9,3	46	8	17,4	48	4	8,3	
SAN_CONSUE	1909	54	2,8	733	22	3	66	9	13,6	95	11	11,6	356	15	4,2	169	11	6,5	258	8	3,1	
SUE_MYSTER	3631	75	2,1	1383	26	1,9	109	8	7,3	330	27	8,2	746	14	1,9	336	12	3,6	265	16	6	
BAL_EUGENI	611	20	3,3	270	12	4,4	13	5	38,5	17	4	23,5	134	10	7,5	14	7	50	87	5	5,7	
BAL_PGORIO	1021	26	2,5	476	15	3,2	28	6	21,4	21	6	28,6	223	9	4	19	8	42,1	151	7	4,6	
BAL_COLCHA	250	13	5,2	119	10	8,4	5	3	60	11	3	27,3	42	7	16,7	14	4	28,6	28	4	14,3	
BAL_LYSVAL	288	17	5,9	106	10	9,4	4	2	50	8	5	62,5	99	8	8,1	22	6	27,3	14	4	28,6	
DUM_MOUSQU	2866	41	1,4	1944	25	1,3	95	12	12,6	112	14	12,5	270	16	5,9	115	11	9,6	95	10	10,5	
BAL_CUREVI	530	23	4,3	235	12	5,1	15	4	26,7	31	10	32,3	96	7	7,3	15	6	40	84	6	7,1	
SUE_JUIFER	4501	85	1,9	1668	25	1,5	164	13	7,9	461	28	6,1	861	12	1,4	474	13	2,7	224	15	6,7	
BAL_TENAFF	671	20	3	346	12	3,5	13	4	30,8	27	5	18,5	96	9	9,4	20	7	35	87	6	6,9	
...

Tabelle B.1: Type-Token-Relation der *verba communicandi*. Nebeneinanderstellung der VC-Types, VC-Tokens und der daraus resultierenden TTR sowohl für die gesamte Redebegleitung als auch für die einzelnen Konstruktionstypen.

Anhang B: Type-Token-Relation der *verba communicandi*

	VC-Tokens ges. (RB)	VC-Types ges. (RB)	TTR RB	T1: VC-Tokens	T1: VC-Types	TTR T1	T2a: VC-Tokens	T2a: VC-Types	TTR T2a	T3a: VC-Tokens	T3a: VC-Types	TTR 3a	T3b: VC-Tokens	T3b: VC-Types	TTR 3b	T3c: VC-Tokens	T3c: VC-Types	TTR 3c	T3e: VC-Tokens	T3e: VC-Types	TTR 3e
DUM_CHRIST	5821	67	1,2	3973	36	0,9	91	14	15,4	226	18	8	615	16	2,6	292	14	4,8	242	15	6,2
SAN_MAREDI	176	10	5,7	106	8	7,5	5	2	40	8	3	37,5	21	4	19	7	2	28,6	17	3	17,6
BAL_COUSPO	919	19	2,1	443	12	2,7	32	7	21,9	64	10	15,6	174	10	5,7	29	5	17,2	75	4	5,3
BAL_COUSBE	1583	23	1,5	726	11	1,5	47	7	14,9	84	11	13,1	350	11	3,1	48	7	14,6	112	8	7,1
DUM_COLLIE	3017	63	2,1	1764	40	2,3	116	16	13,8	254	22	8,7	303	16	5,3	195	16	8,2	91	11	12,1
BAL_PAYSAN	984	30	3	543	15	2,8	23	6	26,1	71	9	12,7	150	12	8	9	6	66,7	67	6	9
BAL_PBOURG	525	15	2,9	278	11	4	15	4	26,7	27	7	25,9	103	9	8,7	10	5	50	48	6	12,5
DUM_TULIPE	652	31	4,8	404	27	6,7	24	11	45,8	28	10	35,7	82	13	15,9	25	7	28	22	9	40,9
SAN_FCHAMP	222	14	6,3	110	10	9,1	5	3	60	13	4	30,8	38	5	13,2	3	2	66,7	31	5	16,1
FLA_BOVARY	622	36	5,8	421	25	5,9	10	5	50	33	13	39,4	81	15	18,5	15	7	46,7	17	4	23,5
FLA_SALAMM	149	23	15,4	115	17	14,8	0	0	0	9	3	33,3	8	5	62,5	1	1	100	6	1	16,7
HUG_MISERA	1526	47	3,1	1254	31	2,5	24	11	45,8	38	12	31,6	33	11	33,3	24	8	33,3	56	8	14,3
GON_GERMIN	126	16	12,7	46	8	17,4	3	1	33,3	6	3	50	13	4	30,8	7	3	42,9	34	5	14,7
VER_TERLUN	371	16	4,3	260	9	3,5	8	4	50	20	7	35	20	6	30	11	5	45,5	9	3	33,3
ZOL_THRAQU	180	21	11,7	83	19	22,9	3	1	33,3	21	6	28,6	16	6	37,5	10	8	80	28	5	17,9
FLA_EDUCAT	532	38	7,1	425	24	5,6	6	6	100	35	7	20	19	6	31,6	8	5	62,5	11	3	27,3
VER_LIEUES	568	26	4,6	465	15	3,2	5	4	80	40	9	22,5	14	6	42,9	4	4	100	11	3	27,3
ZOL_CUREE	307	22	7,2	163	16	9,8	5	4	80	40	13	32,5	25	9	36	21	7	33,3	30	6	20
VER_TOURMO	521	15	2,9	390	14	3,6	10	5	50	43	6	14	28	4	14,3	3	3	100	24	4	16,7
ZOL_VENTRE	370	30	8,1	211	23	10,9	14	6	42,9	43	8	18,6	26	8	30,8	15	9	60	24	5	20,8
ZOL_CONQUE	856	42	4,9	449	31	6,9	28	12	42,9	83	16	19,3	84	15	17,9	33	8	24,2	84	13	15,5
VER_ILEMYS	1708	22	1,3	1423	15	1,1	2	2	100	88	11	12,5	58	7	12,1	7	5	71,4	31	5	16,1
ZOL_EXCELL	631	54	8,6	344	30	8,7	15	9	60	78	16	20,5	49	8	16,3	23	8	34,8	35	10	28,6
ZOL_ASSOMM	587	33	5,6	376	26	6,9	15	8	53,3	50	15	30	40	10	25	9	6	66,7	34	6	17,6
ZOL_PAGEAM	600	24	4	395	20	5,1	11	7	63,6	53	9	17	40	10	25	21	10	47,6	19	4	21,1
ZOL_NANA	830	36	4,3	493	26	5,3	31	8	25,8	103	14	13,6	61	11	18	31	9	29	37	10	27
FLA_BOUVAR	184	24	13	140	17	12,1	4	4	100	10	5	50	1	1	100	1	1	100	4	2	50
ZOL_BONHEU	784	36	4,6	487	29	6	21	10	47,6	93	13	14	37	9	24,3	31	9	29	37	8	21,6
ZOL_GERMIN	514	46	8,9	327	35	10,7	19	12	63,2	43	15	34,9	26	7	26,9	10	7	70	32	9	28,1
ZOL_TERRE	509	43	8,4	332	34	10,2	13	5	38,5	66	18	27,3	15	7	46,7	10	6	60	25	8	32
ZOL_BETEHU	208	30	14,4	133	23	17,3	4	4	100	25	10	40	8	5	62,5	4	3	75	10	3	30
ZOL_ARGENT	246	30	12,2	146	23	15,8	4	4	100	45	14	31,1	11	3	27,3	11	8	72,7	7	4	57,1
ZOL_DEBACL	369	38	10,3	225	27	12	7	5	71,4	49	14	28,6	16	9	56,2	6	6	100	21	8	38,1
PREV_AUTOM	284	27	9,5	196	24	12,2	4	3	75	28	11	39,3	23	9	39,1	9	6	66,7	10	2	20
ZOL_DRPASC	188	25	13,3	110	19	17,3	2	2	100	32	13	40,6	8	3	37,5	5	5	100	5	3	60
PREV_VIERG	386	33	8,5	244	29	11,9	7	4	57,1	27	11	40,7	41	11	26,8	13	7	53,8	20	4	20
ZOL_TRVIL	310	33	10,6	194	23	11,9	7	5	71,4	37	16	43,2	8	3	37,5	4	4	100	22	5	22,7
FOU_GD_ME	128	15	11,7	72	13	18,1	2	1	50	16	7	43,8	9	4	44,4	7	4	57,1	7	3	42,9
PROU_SWANN	273	25	9,2	91	11	12,1	23	7	30,4	20	7	35	29	6	20,7	12	6	50	49	8	16,3
PROU_OMBRE	218	28	12,8	88	12	13,6	3	2	66,7	15	8	53,3	24	6	25	15	4	26,7	30	6	20
COLE_CHERI	227	67	29,5	118	49	41,5	12	8	66,7	39	16	41	8	7	87,5	9	8	88,9	13	4	30,8
PROU_GUERM	584	41	7	240	25	10,4	16	7	43,8	32	12	37,5	62	10	16,1	43	10	23,3	58	6	10,3

Tabelle B.1: Type-Token-Relation der *verba communicandi*. Es werden jeweils die VC-Types, VC-Tokens und die daraus resultierende TTR sowohl für die gesamte Redebegleitung als auch für die einzelnen Konstruktionstypen nebeneinandergestellt. RB = Redebegleitung, TTR = Type-Token-Relation. Die Erläuterung der aufgeführten Konstruktionstypen erfolgt auf S. 113. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

Anhang C

Konstruktionstypen innerhalb des redebegleitenden *discours citant*

	VC-Tokens ges. (RB)	VC-Types ges. (RB)	Anteil T1 (in %)	Anteil T2a (in %)	Anteil T2b (in %)	Anteil T2c (in %)	Anteil T3a (in %)	Anteil T3b (in %)	Anteil T3c (in %)	Anteil T3e (in %)	Anteil Tvoix (in %)	Anteil Toreille (in %)	Anteil TKonj (in %)
VOL_ZADIG	124	12	74,2	0,8	0	0,8	1,6	2,4	0	16,1	0	0	0
VOL_MEMNON	21	6	57,1	9,5	0	0	0	0	0	23,8	0	0	0
VOL_CANDID	270	7	81,1	1,1	0	0	1,9	0,7	0	12,6	0	0	0
VOL_BLANCN	79	8	83,5	1,3	0	1,3	1,3	0	0	7,6	1,3	0	0
VOL_INGENU	61	9	72,1	0	0	0	0	0	1,6	21,3	0	0	0
VOL_PRINCB	60	7	65	0	0	0	0	3,3	0	26,7	0	0	0
SPI_PAUVIR	27	6	70,4	0	0	0	3,7	0	0	22,2	0	0	0
HUG_HANISLA	1136	33	56,4	4,6	0,9	0,4	6,7	7,7	6,2	2,6	6,2	0	0,1
STE_ROUGEN	893	17	41,1	3,4	0,9	0,2	6,5	8,4	5,6	23,6	2,9	0	0,1
BAL_PEAUCH	597	37	44,9	1,5	0,5	1,5	4	23,6	3	9,4	3,9	0	0,2
HUG_NODAME	970	33	60	3,5	0,8	1,1	8	8,7	5,6	3,3	2,3	0	0,1
SND_INDIAN	394	19	40,1	2	0,5	0,5	5,1	17,3	8,1	19,5	3,6	0	0
BAL_MEDCAM	475	37	50,5	0,2	0	0,8	4,2	16,2	3,8	12	2,9	0	0,8
BAL_DUCHES	250	13	35,6	1,2	0	1,2	3,6	29,6	8	11,6	3,2	0	0
BAL_FILLEY	160	12	45	2,5	0	2,5	2,5	20	3,1	14,4	0,6	0	1,9
BAL_CESBIR	906	20	44,8	4,1	0,7	2,9	4,6	18	2,4	12,4	2,9	1	0,2
STE_CHARTR	689	25	35	4,5	1	0,6	3,8	8,3	6,5	29,6	3,8	0	0,3
BAL_FEMMET	448	13	35,5	0,9	0	1,3	7,1	21,4	10,3	10,7	8	0,2	0,9
SND_CONSUE	1909	54	38,4	3,5	0,2	1,5	5	17,3	8,9	13,5	3,5	0,1	0,4
SUE_MYSTER	3631	75	38,1	3	0,2	0,2	9,1	19,3	9,3	7,3	5,6	0	0,1
BAL_EUGENI	611	20	44,2	2,1	0,2	3,1	2,8	21,3	2,3	14,2	3,8	0,2	0,5
BAL_PGORIO	1021	26	46,6	2,7	0,2	0,7	2,1	20,4	1,9	14,8	4,2	0,6	0,2
BAL_COLCHA	250	13	47,6	2	0	0,8	4,4	15,6	5,6	11,2	5,2	0	0,8
BAL_LYSVAL	288	17	36,8	1,4	0	2,8	2,8	33,7	7,6	4,9	4,5	0	0,3
DUM_MOUSQU	2866	41	67,8	3,3	0,6	0,4	3,9	9,1	4	3,3	2,2	0	0,3
BAL_CUREVI	530	23	44,3	2,8	0,2	2,5	5,8	17,7	2,8	15,8	3,4	0,6	0
SUE_JUIFER	4501	85	37,1	3,6	0,4	0,2	10,2	18	10,5	5	7,2	0	0,2
BAL_TENAFF	671	20	51,6	1,9	0	3,9	4	14	3	13	2,5	1	0,4
DUM_CHRIST	5821	67	68,3	1,6	0,6	0,4	3,9	10,2	5	4,2	1,4	0	0,2
...

Tabelle C.1: Verteilung der Konstruktionstypen innerhalb des redebegleitenden *discours citant* für alle im Korpus enthaltenen Texte.

	VC-Tokens ges. (RB)	VC-Types ges. (RB)	Anteil T1 (in %)	Anteil T2a (in %)	Anteil T2b (in %)	Anteil T2c (in %)	Anteil T3a (in %)	Anteil T3b (in %)	Anteil T3c (in %)	Anteil T3e (in %)	Anteil Tvoix (in %)	Anteil Toreille (in %)	Anteil TKonj (in %)
SND_MAREDI	176	10	60,2	2,8	0	1,7	4,5	9,1	4	9,7	1,7	0	0
BAL_COUSPO	919	19	48,2	3,5	0	2,8	7	17,4	3,2	8,2	2,1	1	0,1
BAL_COUSBE	1583	23	45,9	3	0,4	3,6	5,3	21,4	3	7,1	3	1,3	0,2
DUM_COLLIE	3017	63	58,5	3,8	0,7	0,9	8,4	9,4	6,5	3	2,3	0,1	0
BAL_PAYSAN	984	30	55,2	2,3	0	2,9	7,2	14,8	0,9	6,8	2	0,4	0,4
BAL_PBOURG	525	15	53	2,9	0,4	1,7	5,1	18,1	1,9	9,1	1,9	1,3	0,2
DUM_TULIPE	652	31	62	3,7	1,5	0,5	4,3	10,7	3,8	3,4	1,8	0,3	0,3
SND_FCHAMP	222	14	49,5	2,3	0,5	1,8	5,9	17,1	1,4	14	1,4	0	0,5
FLA_BOVARY	622	36	67,7	1,6	0,6	0,3	5,3	12,7	2,4	2,7	1,8	0	0
FLA_SALAMM	149	23	77,2	0	0,7	0	6	4,7	0,7	4	2	0	0
HUG_MISERA	1526	47	82,2	1,6	0,2	0,9	2,5	2,1	1,6	3,7	1,5	0	0,1
GON_GERMIN	126	16	36,5	2,4	0	2,4	4,8	10,3	5,6	27	1,6	0	0,8
VER_TERLUN	371	16	70,1	2,2	0,3	0,3	5,4	4,9	3	2,4	3	0	0
ZOL_THRAQU	180	21	46,1	1,7	0	0,6	11,7	8,3	5,6	15,6	6,1	0	0
FLA_EDUCAT	532	38	79,9	1,1	0	0,4	6,6	3,4	1,5	2,1	0,8	0	0
VER_LIEUES	568	26	81,9	0,9	0,2	0,4	7	2,5	0,7	1,9	1,2	0	0
ZOL_CUREE	307	22	53,1	1,6	0	0,3	13	8,1	6,8	9,8	2,9	0	0
VER_TOURMO	521	15	74,9	1,9	0	0	8,3	5,2	0,6	4,6	2,1	0	0
ZOL_VENTRE	370	30	57	3,8	0,3	1,1	11,6	6,5	4,1	6,5	3,2	0	0,3
ZOL_CONQUE	856	42	52,5	3,3	0,1	0,6	9,7	9,3	3,9	9,8	4,9	0	0
VER_ILEMYS	1708	22	83,3	0,1	0	0	5,2	3,4	0,4	1,8	1,1	0,2	0
ZOL_EXCELL	631	54	54,5	2,4	0	1,1	12,4	7,6	3,6	5,5	4,6	0,6	0
ZOL_ASSOMM	587	33	64,1	2,6	0,2	1,4	8,5	6,6	1,5	5,8	3,4	0,2	0
ZOL_PAGEAM	600	24	65,8	1,8	0	2,3	8,8	6,3	3,5	3,2	2	0	0,2
ZOL_NANA	830	36	59,4	3,7	0	1,4	12,4	7,3	3,7	4,5	2,4	0,2	0,5
FLA_BOUVAR	184	24	76,1	2,2	0	1,1	5,4	0,5	0,5	2,2	1,1	0	0
ZOL_BONHEU	784	36	62,1	2,7	0	1	11,9	4,7	4	4,7	4,3	0	0
ZOL_GERMIN	514	46	63,6	3,7	0	1	8,4	4,7	1,9	6,2	4,1	0,2	0
ZOL_TERRE	509	43	65,2	2,6	0	0,8	13	2,8	2	4,9	4,3	0	0
ZOL_BETEHU	208	30	63,9	1,9	0	0	12	3,8	1,9	4,8	3,8	0	0
ZOL_ARGENT	246	30	59,3	1,6	0,8	0,4	18,3	4,5	4,5	2,8	2,8	0	0
ZOL_DEBACL	369	38	61	1,9	0,5	0,8	13,3	4,3	1,6	5,7	3,8	0	0
PREV_AUTOM	284	27	69	1,4	0	0	9,9	8,1	3,2	3,5	2,5	0,4	0,7
PREV_AUTOM	386	33	63,2	1,8	0,3	1	7	10,6	3,4	5,2	2,1	0,3	0
ZOL_DRPASC	188	25	58,5	1,1	0,5	0,5	17	3,7	2,7	2,7	4,3	0	0
ZOL_TRVIL	310	33	62,6	2,3	0	1,9	11,9	2,3	1,3	7,1	2,6	0	0
AFOU_GD_ME	128	15	56,2	1,6	0	0	12,5	7	5,5	5,5	3,9	0	0
PROU_SWANN	273	25	33,3	8,4	0	4,8	7,3	9,5	4,4	17,9	5,1	0	0,4
PROU_OMBRE	218	28	40,4	1,4	0,5	5	6,9	11	6,9	13,8	2,8	0	0
COLE_CHERI	227	67	52	5,3	0	1,8	17,2	3,1	4	5,7	2,6	0	0
PROU_GUERM	584	41	41,1	2,7	0,7	6,2	5,5	10,3	7,4	9,9	6,5	0	0,2

Tabelle C.1: Verteilung der Konstruktionstypen innerhalb des redebegleitenden *discours citant* für alle im Korpus enthaltenen Texte (Sortierung von oben nach unten chronologisch). Die Angabe der Werte für die Konstruktionstypen erfolgt in Prozent. Bei den Werten für die VC-Types handelt es sich um ganze Zahlen. Die Prozentwerte für die einzelnen Konstruktionstypen ergeben sich aus der Division der Anzahl der VC-Tokens für den jeweiligen Konstruktionstyp durch die Menge aller VC-Tokens innerhalb der Redebegleitung eines Textes, multipliziert mit 100. Die Erläuterung der aufgeführten Konstruktionstypen erfolgt auf S. 113. Die Romankürzel sind nach Tabelle 3.1 aufzulösen.

Danksagung

Ich möchte an dieser Stelle all jenen danken, die wesentlich zum Gelingen dieser Arbeit beigetragen haben. Herrn Prof. Dr. Raible für die Möglichkeit, ein spannendes Thema weiter abseits der linguistisch üblichen Pfade bearbeiten zu können und natürlich für die exzellente und geduldige Betreuung trotz oder gerade wegen der beiden anderen großen „Projekte“ in meinem Leben. Herrn Prof. Dr. Stefan Pfänder für seine Bereitschaft, das Zweitgutachten zu erstellen, und sein stets offenes Ohr. Meiner wunderbaren Familie, die mir auf vielfältige Weise immer wieder den Rücken für die Fertigstellung dieser Arbeit freigehalten hat. Eva Staudinger, Claus Pusch, Joan Ibañez für viele interessante und aufmunternde Gespräche, aus denen sich stets Anregungen und neue Aspekte für die Bearbeitung des Themas ergeben haben. Und all den anderen, deren Unterstützung mir aufgrund des langen Schaffensprozesses vielleicht nicht mehr ganz so präsent ist. Ein besonderer Dank gilt auch meinem geschätzten Lehrer, Herrn Roderich Zeller, ohne den ich womöglich nie eine solche Begeisterung für die Beschäftigung (nicht nur) mit der französischen Sprache entwickelt hätte. Zu guter Letzt möchte ich an dieser Stelle auch die diversen Programme erwähnen, die ganz entscheidend zur inhaltlichen und äußeren Qualität dieser Arbeit beigetragen haben (Perl, TreeTagger, Eclipse und EPIC, gnuplot, yEd, L^AT_EX und einige andere), und deren Entwicklern meinen Dank ausdrücken.

Abbildungsverzeichnis

3.1	DC-Analyse: Programmoberfläche	50
3.2	DC-Analyse: Schematische Darstellung des Programmablaufs	52
3.3	Codebeispiel für das Generieren von konjugierten Verbformen	55
3.4	Beispieldatei: Position des VC	59
3.5	Beispieldatei: eingeleitete Dialogstellen	69
3.6	Beispieldatei: Dialogstellen am Zeilenanfang	70
3.7	Beispieldatei: <i>verba communicandi</i> in redeeinleitenden Sätzen (mit Tags)	71
3.8	Beispieldatei: <i>verba communicandi</i> in redebegleitenden Sätzen (mit Tags)	73
3.9	Beispieldatei: Turns ohne Einleitung	74
3.10	Beispieldatei: Frequenz der <i>verba communicandi</i>	75
3.11	Mangelhafte Textformatierung in <i>Pot-bouille</i> (1882) von Émile Zola.	76
3.12	Mangelhafte Textformatierung in <i>Du côté de chez Swann</i> (1913) von Marcel Proust	77
4.1	Verhältnis von direkter Rede und Erzähltext	85
4.2	Anteil der direkten Rede am Erzähltext	86
4.3	Verteilung von Redeeinleitung, Redebegleitung und alleinstehenden Turns	88
4.4	Verteilung von Redeeinleitung, Redebegleitung und alleinstehenden Turns (in Prozent)	89
4.5	Die häufigsten VC-Types bei Voltaire und Bernardin de Saint-Pierre (absolute und prozentuale Darstellung)	92
4.6	Die häufigsten VC-Types bei Victor Hugo, Stendhal, George Sand und Eugène Sue (absolute und prozentuale Darstellung)	94
4.7	Die häufigsten VC-Types bei Honoré de Balzac (absolute Darstellung).	98
4.8	Die häufigsten VC-Types bei Honoré de Balzac (prozentuale Darstellung).	99
4.9	Die häufigsten VC-Types bei Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert (absolute Darstellung).	102
4.10	Die häufigsten VC-Types bei Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert (prozentuale Darstellung).	103

4.11 Die häufigsten VC-Typen bei Émile Zola (absolute Darstellung).	106
4.12 Die häufigsten VC-Typen bei Émile Zola (prozentuale Darstellung). . . .	107
4.13 Die häufigsten VC-Typen bei Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette (absolute und prozentuale Darstellung)	111
4.14 Übersicht über die Konstruktionstypen (gesamtes Korpus, Redebegleitung)	115
4.15 Verteilung von Konstruktionstyp T1	116
4.16 Verteilung von Konstruktionstyp T2a	117
4.17 Verteilung von Konstruktionstyp T3a	117
4.18 Verteilung von Konstruktionstyp T3b	117
4.19 Verteilung von Konstruktionstyp T3c	118
4.20 Verteilung von Konstruktionstyp T3e	118
4.21 Verteilung der zehn häufigsten VC-Typen pro Konstruktionstyp innerhalb des gesamten Korpus.	129
4.22 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Voltaire und Bernardin de Saint-Pierre.	131
4.23 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Victor Hugo, Stendhal, George Sand und Eugène Sue.	133
4.24 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Honoré de Balzac.	136
4.25 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Alexandre Dumas, Jules Verne, Jules und Édmond de Goncourt und Gustave Flaubert.	137
4.26 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Émile Zola.	140
4.27 Verteilung der häufigsten VC-Typen pro Konstruktionstyp bei Marcel Prévost, Henri Alain-Fournier, Marcel Proust und Colette.	141
4.28 Verhältnis von VC-Tokens und VC-Typen für das gesamte Korpus	145
4.29 Type-Token-Relation der <i>verba communicandi</i> (gesamte Redebegleitung)	147
4.30 Type-Token-Relationen für die Konstruktionstypen T1, T2a und T3a . . .	150
(a) TTR der <i>verba communicandi</i> für Konstruktionstyp T1	150
(b) TTR für Konstruktionstyp T2a	150
(c) TTR für Konstruktionstyp 3a	150
4.31 Type-Token-Relationen für die Konstruktionstypen T3b, T3c und T3e . .	151
(a) TTR für Konstruktionstyp T3b	151
(b) TTR für Konstruktionstyp T3c	151
(c) TTR für Konstruktionstyp T3e	151
4.32 Entwicklung von <i>balbutier</i>	159

4.33 Entwicklung von <i>bégayer</i>	161
4.34 Entwicklung von <i>crier</i>	162
4.35 Entwicklung von <i>s'écrier</i>	163
4.36 Entwicklung von <i>grommeler</i>	164
4.37 Entwicklung von <i>hurler</i>	165
4.38 Entwicklung von <i>murmurer</i>	166
4.39 Entwicklung von <i>interrompre</i>	167
4.40 Entwicklung von <i>soupirer</i>	169
4.41 Entwicklung von <i>éclater</i>	169
4.42 Entwicklung von <i>ricaner</i>	171

Literaturverzeichnis

- Almenberg, Dagmar (1958). Les répliques chez quatre romanciers contemporains. *Studia Neophilologica*. 39: 200–213.
- Authier, Jacqueline (1978). Les formes du discours rapporté. Remarques syntaxiques et sémantiques à partir des traitements proposés. In: *Autour du discours rapporté*. Numéro 17 in Documentation et recherche en linguistique allemande contemporaine. Paris: Université de Paris VIII. 1–87.
- Bachtin, Michail M. (1979). *Die Ästhetik des Wortes*. Frankfurt a. M.: Suhrkamp. 1. Auflage.
- Bachtin, Michail M. & Volosinov, Valentin N. (1977). *Le marxisme et la philosophie du langage: essai d'application de la méthode sociologique en linguistique*. Le sens commun. Paris: Editions de Minuit.
- Bally, Charles (1944). *Linguistique générale et linguistique française*. Bern: Francke. 2. Auflage.
- Bescherelle, Louis Nicolas (1990). *L'art de conjuguer: dictionnaire de 12000 verbes*. Paris: Hatier.
- Bessonnat, Daniel (1994). Paroles des personnages: bâtir une progression. *Pratiques*. 83: 5–34.
- Betten, Anne (1994). Analyse literarischer Dialoge. In: Fritz, Gerd & Hundsnurscher, Franz (Hg.), *Handbuch der Dialoganalyse*. Niemeyer. 519–544.
- Biardzka, Elżbieta (2007). Note sur le verbe du dire en discours direct. Cas de la presse écrite. In: Grabowska, Monika (Hg.), *Le verbe dans tous ses états, Romanica Wratislaviensia*, Band 54. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Bäckvall, Hans (1999). Verbes déclaratifs en français et en espagnol. *Revue de Linguistique Romane*. 63: 509–44.
- Charolles, Michel (1976). Exercices sur les verbes de communication. *Pratiques*. 9: 83–107.
- Chatton, René (1953). *Zur Geschichte der romanischen Verben für ‚sprechen‘, ‚sagen‘*

- und ‚reden‘. Bern: Francke.
- Cicurel, Francine (1994). La conversation sur papier. In: Coleman, James A. & Crawshaw, Robert (Hg.), *Discourse Variety in Contemporary French*. Middlesex: Middlesex University Printing Unit. 121–141.
- Coltier, Danielle (1989). Introduction aux paroles des personnages: fonctions et fonctionnement. *Pratiques*. 64: 69–109.
- Combettes, Bernard (1989). Discours rapporté et énonciation: trois approches différentes. *Pratiques*. 64: 111–122.
- Cornulier, Benoît de (1978). L'incise, la classe des verbes parenthétiques et le signe mimique. *Cahier de linguistique*. 8: 53–95.
- (2004). Sur la valeur de l'incise et sa postposition. In: Benjamins, John (Hg.), *Lexique, syntaxe et lexique-grammaire*, Band 24. Amsterdam: Leclère, Christian. 105–111.
- Cosset, Evelyne (1991). La représentation de «l'acte de parole» des personnages dans *La fortune des Rougon*. *Cahiers naturalistes*. 37 (65): 155–168.
- Delaveau, Annie (1988). La voix et les bruits. Note sur les verbes introducteurs du discours rapporté. *Linx*. 18: 125–135.
- Dessaintes, Maurice (1960). *La construction par insertion incidente*. Paris: Éditions d'Artrey.
- Durrer, Sylvie (1994). *Le dialogue romanesque. Style et structure*. Genève: Droz.
- Éjchenbaum, Boris (1969). Die Illusion des 'skaz. In: Striedter, Jurij & Stempel, Wolf-Dieter: (Hg.), *Texte russischer Formalisten. Band 1: Texte zur Literaturtheorie und zur Theorie der Prosa.*, Band 1. München: Fink.
- Fairon, Cédric (2004). Une étude de corpus pour éclairer la question du verbe de l'incise en français. In: Leclère, Christian (Hg.), *Lexique, syntaxe et lexique-grammaire*. *Linguisticae investigationes: Supplementa = Studies in French & general linguistics*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Fónagy, Ivan (1984). Reported speech in French and Hungarian. In: Coulmas, Florian (Hg.), *Direct and indirect speech*. Mouton de Gruyter. 254–309.
- Gallèpe, Thierry (2002). Redewiedergabe: ein paradoxer Begriff. In: Baudot, Daniel (Hg.), *Redewiedergabe, Redeerwähnung. Formen und Funktionen des Zitierens und Reformulierens im Text*. Tübingen: Stauffenburg-Verlag. 57–67.
- Gaulmyn, Marie-Madeleine de (1986). Sur les verbes de parole. Dire, demander, proposer. In: *Sur le verbe*. Lyon: Presses universitaires de Lyon.

- Głowinski, Michał (1974). Der Dialog im Roman. *Poetica*. 6: 1–16.
- Goetsch, Paul (1985). Fingierte Mündlichkeit in der Erzählkunst entwickelter Schriftkulturen. *Poetica*. 17 (3-4): 202–218.
- Gothot-Mersch, Claudine (1969). Le dialogue dans l'œuvre de Flaubert. *Europe*. 485, 486, 487: 112–128.
- (1981). De Madame Bovary à Bouvard et Pécuchet: la parole des personnages dans le roman de Flaubert. *Revue d'histoire littéraire de la France*. 81: 542–562.
- Gougenheim, Georges (1938). La présentation du discours direct dans «La Princesse de Clèves» et dans «Dominique». *Le français moderne*. 6: 305–320.
- Grevisse, Maurice & Goosse, André (1993). *Le bon usage*. Paris: Duculot.
- (2008). *Le bon usage*. Bruxelles: De Boeck, Duculot.
- Grosse, Siegfried (1972). Literarischer Dialog und gesprochene Sprache. In: Backes, Herbert (Hg.), *Festschrift f. Hans Eggers zum 65. Geburtstag*. Tübingen: Niemeyer. 649–669.
- Gülich, Elisabeth (1978). Redewiedergabe im Französischen. Beschreibungsmöglichkeiten im Rahmen einer Sprechakttheorie. In: Meyer-Hermann, Reinhard (Hg.), *Sprechen, Handeln, Interaktion, Konzepte der Sprach- und Literaturwissenschaft*, Band 26. Tübingen: Niemeyer.
- (1990). Erzählte Gespräche in Marcel Prousts *Un amour de Swann*. *Zeitschrift für französische Sprache und Literatur*. 100: 89–109.
- Günther, Werner (1928). *Probleme der Rededarstellung: Untersuchungen zur direkten, indirekten und ‚erlebten‘ Rede im Deutschen, Französischen und Italienischen*. Gräfenhainichen: Schulze.
- Haig, Stirling (1986). *Flaubert and the gift of speech: dialogue and discourse in four 'modern' novels*. Cambridge: Cambridge University Press.
- Halmøy, Odile (1983). La Structure des propositions en incise de type 'dit-il'. *Moderna Språk*. 77 (3): 233–246.
- Henry, Albert (1961). L'expressivité du dialogue dans le roman. In: *La littérature narrative d'imagination*. Paris: Presses Universitaires de France. 3–22.
- Heuvel, Pierre van den (1978). Le discours rapporté. *Neophilologus*. 62: 19–38.
- Hoberg, Susanne (2009). Fingierte Mündlichkeit in der französischen Romanliteratur zwischen 1750 und 1850. Eine computergestützte Analyse. In: Kailuweit, Rolf & Pfänder, Stefan (Hg.), *Franko-Media: Aufriss einer französischen Sprach- und Medienwissenschaft*, Band 17. Berlin: Berliner Wissenschafts-Verlag. Band 17. 99–108.

- Jamrozik, Elżbieta (1992). *La syntaxe et la sémantique des verbes de parole français*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Kalverkämper, Hartwig (1991). Literatur und Körpersprache. *Poetica*. 23 (3/4): 328–373.
- (1997). Körpersprache in älteren Kulturen der Romania. In: Huberty, Maren & Perlich, Claudia (Hg.), *Studia historica romanica. In honorem Johannes Klare*. Bonn: Romanistischer Verlag. 215–243.
- Kerbrat-Orecchioni, Catherine (1996). Dialogue romanesque et conversations naturelles. In: Cmejrková, Svetla (Hg.), *Dialoganalyse VI: Referate der 6. Arbeitstagung. Prag 1996*. Tübingen: Niemeyer. 331–343.
- Kurz, Josef (1966). *Die Redewiedergabe. Methoden und Möglichkeiten*. Leipzig: Karl-Marx-Universität.
- Körner, Karl-Hermann (1977). *Einführung in das semantische Studium des Französischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Lamiroy, Beatrice & Charolles, Michel (2008). Les verbes de parole et la question de l'(in)transitivité. *Discours*. 2. URL <http://discours.revues.org/3232>.
- Lane-Mercier, Gillian (1989). *La parole romanesque*. Ottawa: Presses de l'Université d'Ottawa.
- Lüdeling, Anke (2007). Das Zusammenspiel von qualitativen und quantitativen Methoden in der Korpuslinguistik. In: Kallmeyer, Werner & Zifonun, Gisela (Hg.), *Sprachkorpora: Datenmengen und Erkenntnisfortschritt*. Berlin: de Gruyter. 28–48.
- Mangueneau, Dominique (1981). *Approche de l'énonciation en linguistique française*. Paris: Hachette.
- (1986). *Éléments de linguistique pour le texte littéraire*. Paris: Bordas.
- Marczak, Ewelina (2007). Dire, hasarder les demi-mots, ouvrir la bouche... Quelques remarques sur les verbes de communication dans *Armance* de Stendhal. In: Grabowska, Monika (Hg.), *Le verbe dans tous ses états, Romanica Wratislaviensia*, Band 54. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Marnette, Sophie (2005). *Speech and Thought Presentation in French: Concepts and Strategies*. Band 133. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Martins-Baltar, M. (1976). Les verbes transpositeurs du discours rapporté. In: Gauvenet, Hélène (Hg.), *Pédagogie du discours rapporté*. Paris: Didier.
- Mathet, Marie-Thérèse (1988). *Le dialogue romanesque chez Flaubert*. Paris: Diffusion aux amateurs de livres.

- Messelken, Hans (1996). Computerunterstützte Analysen textsortenspezifischer Lexik. In: Kalverkämper, Hartwig & Baumann, Klaus-Dieter (Hg.), *Fachliche Textsorten: Komponenten, Relationen, Strategien*. Tübingen: Gunter Narr Verlag.
- Monville-Burston, Monique (1993). Les Verba dicendi dans la presse d'information. *Langue Francaise*. 98: 48–66.
- Müller, Andreas (1981). *Figurenrede. Grundzüge der Rededarstellung im Roman*. Dissertation. Georg-August-Universität zu Göttingen. Göttingen.
- Mylne, Vivienne G. (1987). The Uses of Dialogue in Hugo's Early Fiction. *French Studies: A Quartely Review*. 41 (4): 408–420.
- (1994). *Le Dialogue dans le roman français de Sorel à Sarraute*. Paris: Universitas.
- Pérennec, Marie-Hélène (2002). Redewiedergabe in fiktiven und nicht-fiktiven Texten. In: Baudot, Daniel (Hg.), *Redewiedergabe, Redeerwähnung. Formen und Funktionen des Zitierens und Redformulierens im Text*. Tübingen: Stauffenburg. 41–53.
- Prince, Gérald (1978). Le discours attributif et le récit. *Poétique*. 35: 305–313.
- (1980a). Le discours attributif dans «La Peste». In: Gay-Crosier, Raymond (Hg.), *Albert Camus 1980*. Gainesville: University Presses of Florida.
- (1980b). On Attributive Discourse in «Madame Bovary». In: Mitchell, Robert L. (Hg.), *Pretext, text, context. Essays on Nineteenth-Century French Literature*. Columbus: Ohio State University Press.
- Rey-Debove, Josette & Rey, Alain (1993). *Le Nouveau Petit Robert*. Paris: Dictionnaires Le Robert.
- Riegel, Martin/ Pellat, Jean-Christophe & Rioul, René (2009). *Grammaire méthodique du français*. Quadrige : Manuels. Paris: Presses Univ. de France. 4. Auflage.
- Rosier, Laurence (1999). *Le discours rapporté. Histoire, théories, pratiques*. Paris, Bruxelles: Duculot.
- (2008). *Le discours rapporté en français*. Paris: Ophrys.
- Rousset, Jean (1998). *Dernier regard sur le baroque*. Paris: Corti.
- Rullier-Theuret, Françoise (2001). *Le dialogue dans le roman*. Paris: Hachette.
- Sabban, Annette (1988). Verben der Redeeinleitung im Französischen und Deutschen. Ein Beitrag zum Problem der Redeerwähnung. *Zeitschrift für französische Sprache und Literatur*. 88: 28–63.
- Stefani-Meyer, Georgette (2003). Die Redewiedergabe im Französischen. In: Gil, Alberto & Schmitt, Christian (Hg.), *Aufgaben und Perspektiven der romanischen Sprachgeschichte im dritten Jahrtausend*. Bonn: Romanistischer Verlag.

- Suwala, Halina (1987). Le discours attributif dans «La curée». In: Baguley, D. (Hg.), «*La curée*» de Zola ou «*la vie en outrance*». *Actes du colloque du 10 janvier 1987*. Paris: Sedes.
- Ullmann, Stephen (1964). *Style in the french novel*. Oxford: Blackwell. 2. Auflage.
- Vinogradov, Viktor (1969). Das Problem des 'skaz' in der Stilistik. In: Striedter, Jurij & Stempel, Wolf-Dieter (Hg.), *Texte russischer Formalisten. Band 1: Texte zur Literaturtheorie und zur Theorie der Prosa*, Band 1. München: Fink. 169–207.
- Vivès, Robert (1998). Les mots pour le dire: vers la constitution d'une classe de prédicats. *Langages*. 131: 64–76.
- Vogt, Jochen (1998). *Aspekte erzählender Prosa: eine Einführung in Erzähltechnik und Romantheorie*. Opladen: Westdeutscher Verlag. 8. Auflage.
- Wierzbicka, Anna (1974). The semantics of direct and indirect discourse. *Papers in Linguistics*. 7 (3-4): 267–307. URL <http://www.tandfonline.com/doi/abs/10.1080/08351817409370375>.

Während der *discours cité/direct*, d.h. die direkte Rede, in der französischen Romanliteratur Gegenstand zahlreicher Überlegungen und Untersuchungen war und immer noch ist, wurde dem *discours citant* (auch als Redebegleitung bzw. Redeeinleitung bezeichnet), zu keinem Zeitpunkt eine ähnlich große Aufmerksamkeit zuteil. Obwohl letzterer regelmäßiger Bestandteil der Figurenrede ist und durch seine metasprachliche Funktion wesentlich zur Charakterisierung der sprechenden Figuren beiträgt, gibt es nur wenige Einzelbetrachtungen dazu, die obendrein rein qualitativer Natur sind.

Auf der Grundlage eines elektronischen Korpus (knapp 10 Mio. Tokens) und eines eigens für diesen Zweck entwickelten Software-Tools, DCAnalyse, führt die Autorin eine sowohl quantitative als auch qualitative Untersuchung zum *discours citant* durch. Neben einem einführenden theoretischen Teil sowie Erläuterungen zum Aufbau und zur Funktionsweise der Software erfolgt eine Analyse vor allem des redebegleitenden *discours citant* in 70 literarischen Texten von 16 Autoren, deren Werke zwischen 1747 und 1922 entstanden sind. Der Schwerpunkt liegt dabei auf der Analyse der *verba communicandi* sowie häufig in der Redebegleitung wiederkehrender Konstruktionstypen. Ein weiteres Hauptaugenmerk ist die Erörterung der Frage, in wieweit die redebegleitende Konstruktion mit *gérondif* – z.B. "...", *dit-il en murmurant*. – eine Refunktionalisierung erfahren hat, die dazu führt, dass das Verb aus der *gérondif*-Form ab einem bestimmten Zeitpunkt vor allem als vollwertiges Redeverb – z.B. "...", *murmura-t-il*. – eingesetzt wird.

Aus den verschiedenen Fragestellungen ergibt sich erstmalig ein differenziertes Bild der diachronen Entwicklung vor allem der Redebegleitung, das den sich allmählich vollziehenden Wandel der Diskurstradition widerspiegelt.

ISBN 978-3-928969-15-4



9 783928 969154

UNI
FREIBURG