

## Quantitative Analyse

### Vorbereitungen

Zum Verkoden des Textes wurde ATLAS/ti 1.0 (Muhr 1993) eingesetzt. Gegenüber anderen Softwares, die zu diesem Zweck entwickelt wurden (Weitzman und Miles 1995) weist ATLAS/ti mehrere Vorteile auf: Das Programm ermöglicht unter einer grafischen Benutzeroberfläche sowohl maschinelle wie manuelle Verkodung. Es schreibt auf Wunsch kodierte Textstellen und Belege, nach Kategorien sortiert, in Dateien, die von jeder Textverarbeitung gelesen werden können. Ferner ist die Erstellung eines SPSS-kompatiblen Datensatzes möglich. Der einzulesende Textkorpus muß im ASCII-Format vorliegen. Analyseeinheiten werden durch eine Leerzeile getrennt. Insgesamt brachte der Einsatz ATLAS/ti erhebliche Arbeitserleichterung und Zeitersparnis<sup>28</sup>.

### Hardwarebedarf

Die Arbeit mit ATLAS/ti zieht neben den oben genannten Vorteilen allerdings auch extreme Anforderungen an die benutzte Hardware nach sich. Diese sind weitaus höher, als vom Hersteller angegeben. Keinesfalls sollte das "Herumspielen" mit den mitgelieferten Demonstrationstexten zu voreiligen Schlüssen im Hinblick auf den Ernstfall verleiten. Ein 486 DX2-66 mit 12 Mb Arbeitsspeicher war die in der Praxis benötigte Mindestausstattung. 16 Mb und ein Pentium-Prozessor wären in höchstem Maße wünschenswert gewesen.

Unter ATLAS/ti wird eine sogenannte *hermeneutische Einheit* (HU) erzeugt, eine Datei, die Angaben über den benutzten Datensatz und die vom Forscher vorgenommenen Manipulationen enthält. Dazu gehören die Rohdaten, also etwa die zu analysierenden Interviews, die als *Primärtexte* bezeichnet werden. Hinzu kommt eine Darstellung des benutzten Kategorienschemas, in Form einer *Kodelliste*. Anmerkungen zum Text, sogenannte *Memos*, werden auf der Bildschirmfläche dargestellt und intern mit den Rohdaten verknüpft. Beim Kodieren wird dann ein Kode mit einer vom Forscher definierten Passage des Textes verbunden. Angaben über diese Verbindung, eine Darstellung der ersten Wörter des betreffenden Zitats sowie Angaben über die Position des Zitats im laufenden Text werden in einem besonderen *Fenster* der HU dargestellt. Bei größeren Textmengen

---

28. Man vergleiche Giegler (1992), um einen Eindruck von der stürmischen Entwicklung auf diesem Gebiet zu gewinnen.

entsteht so eine Arbeitsdatei, deren Aufruf in den Arbeitsspeicher bei meiner Hardwarekonfiguration minutenlange Wartezeiten mit sich brachte. Gleiches gilt für das Erzeugen von Ausdrucken von Kode- und Zitatlisten sowie SPSS-Outputdateien.

### Besonderheiten

Eine Schwierigkeit tritt bei der Bearbeitung von Gruppendiskussionen auf. Unter ATLAS/ti wird jedes Interview als separate Textdatei in die HU geladen. Während also im Normalfall eine Textdatei einem Informanten entspricht, gilt dies für die Ausgangsdaten einer Gruppendiskussion nicht. Da diese auch nicht auseinandergerissen werden sollten, um die segmentübergreifenden Kontexte abzufangen, mußte also jedes Segment prinzipiell mit einem Kode für den entsprechenden Informanten belegt werden. Hinzu kamen die inhaltlichen Kodierungen. Beim SPSS-Output erstellt ATLAS/ti eine Matrix, auf der die kodierten Segmente die Fälle, die Kodes als Variablen abgebildet werden. Diese Matrix gibt dann allerdings keine Auskunft über die individuelle Verknüpfung von Informanten und Kodes. Die Kodes, die für die Informanten vergeben wurden, werden als weitere Variablen dargestellt. Dieses Problem war unter SPSS nur durch *brute force* zu überwinden, in dem Kreuztabellen aller Variablen, die Angaben über die Informanten enthielten, mit denen, die die Kodes abbildeten, durchgeführt wurden, und die entsprechenden Werte in eine neue Matrix (Informanten als Fälle, inhaltliche Kodes als Variablen) übertragen wurden. Verständlicherweise ist dies ab einer größeren Anzahl von Fällen, sowohl was die Anzahl der Informanten und Kategorien, wie auch des erzeugten Korpus betrifft, nicht mehr praktikabel.

Um diese Arbeit in Zukunft zu erleichtern, hat Gerd Jellinghaus<sup>29</sup> ein Programm geschrieben, das diesen Vorgang automatisiert. Mit Hilfe dieser Utility können zwei verschiedene Variablenbereiche in einer Matrix definiert werden. Danach wird für jeden Fall und für jede Variable in beiden Bereichen geprüft, ob das Auftreten eines Wertes im ersten mit einem aus dem zweiten Variablenbereich einhergeht. Der erste Bereich stellt die Fälle, der zweite die Variablen einer neu zu erzeugenden Matrix dar. Die entsprechenden Werte werden addiert und ihre Summe als Werte im neuen Variablenbereich eingetragen. Der im ASCII-Format vorliegende Datensatz kann in jede Statistiksoftware eingelesen werden. Per

---

29. Der Autor ist unter [a1713272@smail.rz.uni-koeln.de](mailto:a1713272@smail.rz.uni-koeln.de) auf dem INTERNET zu erreichen. Die Software kann auch über den Verfasser ([Oliver.Kortendick@rz.uni-koeln.de](mailto:Oliver.Kortendick@rz.uni-koeln.de)) bezogen werden.

Voreinstellung wird ein ANTHROPAC<sup>30</sup>-konformer Datendeklarationsheader eingesetzt.

## Beschreibung des Datensatzes

### Kategorienauslastung

Abbildung 6 zeigt, in welchem Maße das benutzte Kategoriensystem "griff". Die Angaben in Prozent wurden gerundet. Mit "Griffigkeit des Kategoriensystems" ist hier das Ausmaß gemeint, in dem die zur Verfügung stehenden Kategorien im Verlauf des Verkodungsprozesses eingesetzt wurden.

Die Grafik zeigt, daß bereits nach einem Drittel der zu verkodenden

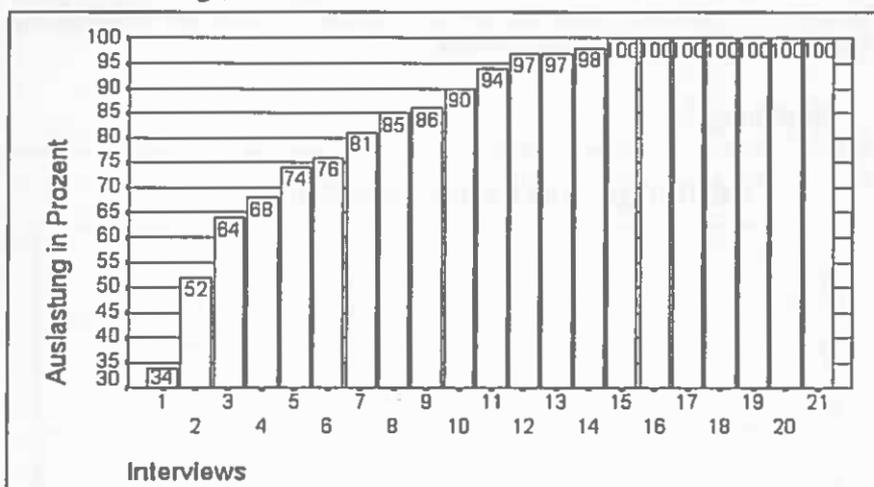


Abbildung 6 Kategorienauslastung

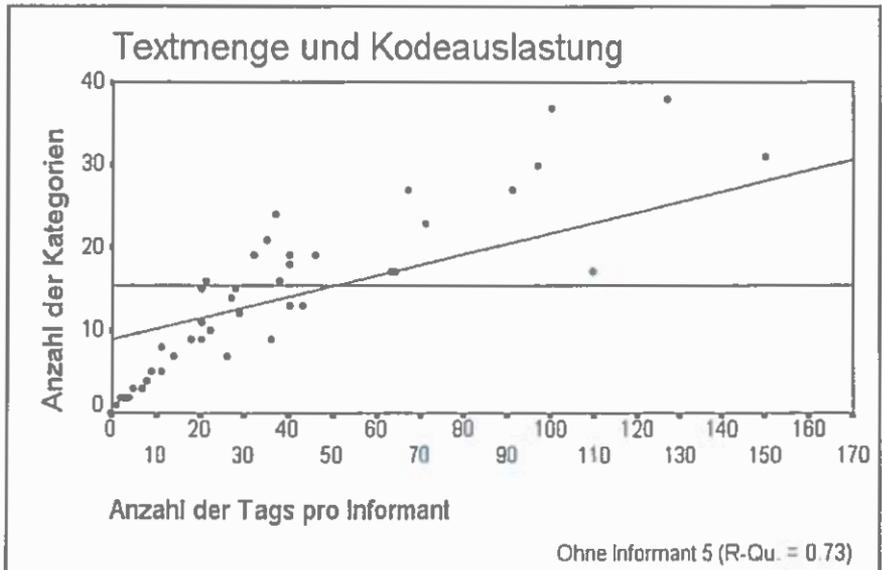
Textmenge 80% der Kategorien gefallen sind. Nach 2 Dritteln erhöht sich dieser Anteil auf 98%. Dies läßt mehrere Schlüsse zu. Es zeigt sich, daß sich das Kategoriensystem, in Bezug auf die Erfassung der wichtigsten Inhalte des Textes, trotz aller Überarbeitung nicht allzusehr von der ursprünglichen Themenliste entfernt hat. Es muß dabei aber berücksichtigt werden, daß sowohl bei der Entwicklung der Themenliste, als auch bei der endgültigen Verkodung des Textes anhand des Kategorienschemas die Interviews in der gleichen Reihenfolge vorlagen. Letztendlich kann die

Griffigkeit des Systems nur dann beurteilt werden, wenn diese Abfolge geändert würde, oder neues Material vorläge.

In welchem Verhältnis sich Segmente, benutzte Kategorien und produzierte Textmenge im Hinblick auf die Informanten verteilen, ergibt sich aus Tabelle 7.

Hinter dem Informantenetikett ist in Klammern die Fallnummer angegeben. "N Abs" bezeichnet die Gesamtzahl der vergebenen inhaltlichen Codes, wobei manche Segmente mit mehreren Codes belegt sein können. "N Kat" gibt die Anzahl der vom jeweiligen Informanten benutzten Kategorien wieder, "N Tags" die insgesamt produzierten Segmente des jeweiligen Informanten, gibt also Auskunft an seinem Anteil im Korpus. Pro Segment wurden im Durchschnitt 3.3 Codes vergeben. Textmenge und die Anzahl der verschiedenen Kategorien stehen in einem linear positiven Verhältnis, wie Abbildung 7 darstellt.

Abbildung 7



Informant Fünf, der extrem vom gesamten Sample abwich, wurde nicht abgebildet, um die Grafik übersichtlich darstellen zu können. Mit 63 verschiedenen Kategorien bei 518 insgesamt vergebenen Codes paßt er sich jedoch dem Trend an. Diese Angaben wurden aber sowohl bei der Mittelwertberechnung (15,3) wie der Regressionsgeraden berücksichtigt.

Korrelationsberechnungen zwischen den drei in Tabelle 7 dargestellten Variablen ergeben einen signifikant hohen *Pearsons r*. Dies bedeutet, daß Textmenge, benutzte Kategorien und gefilterter Redeanteil über alle Informanten betrachtet, in einem engen Zusammenhang stehen:

**Tabelle 6: Pearsons 'r' für den Zusammenhang zwischen Kodes und Tags**

	- - Correlation Coefficients - -	
	N abs	N Kat
N Kat	,8527	
	( 43)	
	P= ,000	
N Tags	,9644	,7258
	( 43)	( 43)
	P= ,000	P= ,000

(Coefficient / (Cases) / 2-tailed Significance)  
 " . " is printed if a coefficient cannot be computed

Diese überzeugend starken Zusammenhänge zeigen unmittelbar die Effizienz der Gruppendiskussion als Erhebungsinstrument. Es ist keinesfalls so, daß bestimmte Informanten nur auf wenige Themen beschränkt blieben und diese öfter wiederholten. Steigt der Redeanteil, so steigt auch die Anzahl der benutzten Themen. Hätten alle Informanten die gleiche Chance eine bestimmte Textmenge zu produzieren, so würden sie im Verlauf ihrer Ausführungen auch nahezu alle Themen ansprechen. Es könnte mit der geringen externen Steuerung durch den Moderator zusammenhängen, daß dies so ungleichmäßig passiert ist. Wird die Diskussion also stärker *gelenkt*, einzelne Teilnehmer entweder zu mehr Beiträgen motiviert oder aber stärker gebremst, so wächst zwar der Anteil von Trübungen durch Interviewerinterventionen, insgesamt aber würden wahrscheinlich homogenere Textanteile der einzelnen Informanten entstehen. Dadurch könnten einige der Schwierigkeiten bei den nun folgenden Auswertungen vermieden werden.

### Darstellung der auftretenden Themen

Angesichts der höchst unterschiedlich produzierten Textmengen erschien es nicht sinnvoll, mit den absolut aufgetretenen Themenhäufigkeiten zu arbeiten. Im folgenden interessiert weniger, wie oft ein bestimmtes Thema (eine Kategorie) im Korpus von einem Informanten benutzt wird, sondern vielmehr, ob es überhaupt vorkommt. Die *Informant-Thema-Matrix* wurde also dichotomisiert, d.h. in Werte zwischen 0 (Thema tritt nicht auf) und 1 (Thema tritt auf) umgewandelt. Für die folgenden Analysen wurden außerdem aus Gründen der Übersichtlichkeit die Themen zu Gruppen zusammengefaßt.

**Tabelle 7: Verteilung von Kategorien und Tags auf die Informanten**

Informanten	N abs	N Kat	N Tags
INF01 (1)	46	19	183
INF02 (2)	11	5	54
INF03 (3)	5	3	32
INF04 (4)	18	9	137
INF05 (5)	518	63	2608
INF06 (6)	38	16	151
INF07 (7)	64	17	176
INF08 (8)	40	13	132
INF09 (9)	91	27	187
INF10 (10)	40	19	27
INF11 (11)	8	4	15
INF12 (12)	3	2	3
INF13 (13)	4	2	6
INF15 (14)	110	17	434
INF16 (15)	22	10	22
INF17 (16)	28	15	116
INF18 (17)	29	12	120
INF19 (18)	11	8	62
INF20 (19)	2	2	5
INF21 (20)	97	30	482
INF22 (21)	35	21	37
INF23 (22)	100	37	131
INF24 (23)	37	24	24
INF25 (24)	0	0	5
INF26 (25)	67	27	125
INF27 (26)	71	23	110
INF28 (27)	36	9	155
INF29 (28)	150	31	298

INF30 (29)	26	7	86
INF31 (30)	127	38	357
INF32 (31)	20	11	82
INF33 (32)	14	7	34
INF34 (33)	9	5	19
INF35 (34)	1	1	6
INF36 (35)	21	16	38
INF37 (36)	32	19	32
INF38 (37)	20	15	19
INF39 (38)	27	14	39
INF40 (39)	43	13	120
INF41 (40)	20	9	43
INF42 (41)	40	18	67
INF44 (42)	7	3	23
INF46 (43)	63	17	319
TOTAL 43	2.151	658	7.121

Mittelwert                      50,02                      15,30                      165,60

Std. Deviation	81,36	12,16	398,47
Std. Error	12,40	1,85	60,77
Median	29	14	67
Mode	20	2	5
Varianz	6.619,69	147,98	15.8781,77

**Tabelle 8: Zusammenfassung der Kategorien in Gruppen**

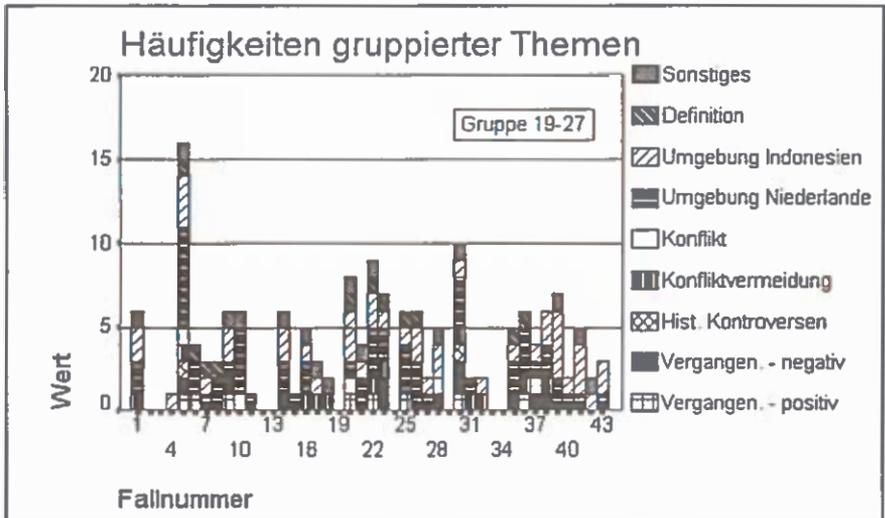
Gruppe	Kategoriennummer	Gruppe	Kategoriennummer
Kindheit positiv	1.1.1. - 1.1.3.	Kindheit negativ	1.2.1. - 1.2.2.
Sozialisation	1.3.1. - 1.3.2.	Distanz	1.4.1. - 1.4.3.
Soziale Beziehungen	1.5.1. - 1.5.2.	Frauen - Kultur	1.6.1.
Negative Erfahrungen	1.7.1. - 1.7.6.	Anekdoten	2.1.1. - 2.1.21.
Indische Identität	3.1.1.	Niederländische Identität	3.2.1.
Ambivalente Identität	3.3.1. - 3.3.2.	Eigenschaften INL	3.4.1. - 3.4.13.
Eigenschaften NL	3.5.1. - 3.5.2.	Transmission	3.6.1.
Reise negativ	4.1.1. u. 4.1.3. u. 4.1.5.	Reise positiv	4.1.2. u. 4.1.4.
Vergangenheit (Bez.)	5.1.1.	Vergangenheit (o.B.)	5.1.2.
Positiver Bezug	5.2.1. - 5.2.2.	Negativer Bezug	5.3.1. - 5.3.3.
Historische Kontroversen	5.4.1.	Konfliktvermeidung	6.1.1.
Konflikt	6.1.2.	Umgebung Niederlande	7.1.1. - 7.1.7.
Umgebung Indonesien	7.2.1. - 7.2.4.	Definition	8.1.
Sonstiges	8.2.	Total:	27

Die Abbildungen 8, 9 und 10 geben Auskunft über die Verteilung dieser Themengruppen für jeden einzelnen Informanten. Es steht jeweils ein Teilstrich für ein Thema innerhalb einer Gruppe. Die Form dieser Darstellung bereinigt gewissermaßen die unterschiedlich großen Textanteile. Trotzdem ist es wohl nicht möglich, die Daten zum Vergleich der Informanten *untereinander* heranzuziehen. Betrachtet man zum Beispiel die Gruppe "Anekdoten", so ist sie zwar unterschiedlich stark ausgeprägt, dies könnte aber immer noch auf die unterschiedlichen Redeanteile zurückzuführen sein, weil bei denjenigen, die extrem viel reden, die Chance höher ist, viele Themen anzusprechen. Bei Informant Fünf spielt außerdem die Anwesenheit bei vielen Diskussionen eine Rolle. Die Abbildungen geben vielmehr Auskunft über die Verteilung von Themengruppen am gesamten Redeanteil eines einzelnen Informanten. Sie spiegeln die Struktur der biographischen Darstellung und die relative Wichtigkeit einzelner Themengruppen für den jeweiligen Informanten wider. Betrachtet man beispielsweise Informant Fünf (Fall 5) kann man feststellen, daß dessen Kenntnis des Anekdotenschatzes fast vollständig ist:

Von 21 hier möglichen Themen sind 17 Bestandteil seiner Ausführungen. Die Gruppe "negative Erfahrungen" ist zu 100% abgedeckt. Dagegen etwa Informant 10 (Fall 10): Er benutzt nur 19 der zur Verfügung stehenden Kategorien. Sein Anteil an den anekdotischen Elementen ist mit drei Ausprägungen relativ gering. Von den 6 vorhandenen Themen zu "negativen Erfahrungen" benutzt er jedoch fünf, um seinen Lebenslauf darzustellen.



Abbildung 10



Die abgebildete Struktur wirft mehrere Fragen auf: Zum einen nach möglichen Zusammenhängen in der Darstellung einzelner Informanten mit anderen, externen Variablen wie *Alter* und *Geschlecht*; andererseits gilt es zu prüfen, ob es einen meßbaren Zusammenhang zwischen Themen gibt, mit anderen Worten, ob die Erwähnung eines bestimmten Themas, die Nennung eines anderen nach sich zieht. Deutlich ist abgebildet, daß es Schwerpunkte in der Darstellung gibt, was einzelne Themenfelder betrifft. So sind die *Anekdoten* überaus stark vertreten, ebenso wie die *Eigenschaften Indischer Niederlander* und *Umgebung Niederlande*. Dies ist sicherlich nicht allein damit erklärt, daß es auch mehr Ausprägungen im Kategoriensystem gibt. Das Kategoriensystem spiegelt die Struktur des gesamten Korpus wider. Keinesfalls kann hieraus aber auf die Erzählstruktur einzelner Informanten geschlossen werden. So ist es überraschend, daß sich die Struktur des Kollektivs unmittelbar auf das Individuum übertragen läßt. Es gibt also keine ausgesprochenen Experten für die Wiedergabe dieser drei Bereiche. Die Frage der Zusammenhänge von Themengruppen und externen Variablen wird im folgenden Abschnitt, der Zusammenhang von Themen im Kapitel "Multivariate Analysen" untersucht. Schließlich interessiert dann noch der Zusammenhang von Informanten aufgrund benutzter Themen.

## Zusammenhänge

Im nächsten Auswertungsschritt sollen Zusammenhänge zwischen den verwendeten Themen und externen Variablen ermittelt werden. Ich beschränke mich in diesem Zusammenhang auf die folgenden drei Merkmale: Alter, Geschlecht und Länge des Aufenthaltes in Indonesien.

Hinsichtlich des Alters der Informanten wurde eine Einteilung in zwei Gruppen vorgenommen. Die Generationszugehörigkeit wurde ausgehend von *ego* beschrieben. Familienmitglieder, die lateral zu *ego* stehen, sind "Die mittlere Generation", Mitglieder, die sich linear unter *ego* befinden, "Die Jüngeren", oberhalb von *ego* "Die Älteren". Diese und die mittlere Generation wurden zusammengefaßt.

Mit Aufenthalt in Indonesien sind Lebensabschnitte gemeint, die einen längeren Zeitraum überschreiten. Die Grenze wurde hier bei Informant 21 mit zwei Jahren gezogen. Die anderen Informanten waren entweder noch länger in Indonesien, oder erheblich kürzer, d.h sie verbrachten allenfalls Urlaube von mehreren Wochen dort.

Die Zusammenhänge wurden durch Kreuztabellierung der dichotomisierten, nach Tabelle 8 gruppierten Variablen ermittelt. Auch diese wurden wiederum dichotomisiert, also auf die Aussage beschränkt, "Gruppe tritt auf oder nicht". Aufgrund der geringen Fallzahl mußte dieser Informationsverlust in Kauf genommen werden. Im folgenden werden die Fälle beschrieben, in denen signifikante Zusammenhänge ermittelt werden konnten, und die vor dem ethnographischen Hintergrund sinnvoll zu interpretieren waren.

Von den 26 Themengruppen (Tabelle 8) wurde mit Hilfe der drei externen Variablen insgesamt 78 mal auf Zusammenhang geprüft. Die letzte Gruppe "Sonstiges" wurde im folgenden nicht berücksichtigt, da hier keine sinnvollen Interpretationen möglich sind. Tabelle 9 zeigt zunächst die Fälle (61), wo keine Zusammenhänge festgestellt werden konnten. Interessant sind diese fehlenden Zusammenhänge, weil sie meiner Meinung nach Themengruppen darstellen, die unabhängig von den externen Variablen von den Informanten angesprochen werden. Das scheint mir deswegen wichtig, weil es ein Hinweis darauf ist, daß diese Inhalte die "stabilsten" im Übertragungsprozeß sind, insbesondere dort, wo sich bei allen drei externen Merkmalen kein Zusammenhang findet. So etwa bei den Gruppen "Kindheit negativ" und "Positiver Bezug zur Vergangenheit". Die damit verbundenen Inhalte werden von den Informanten gleichermaßen in den Gesprächen erwähnt, sind somit zum Bestandteil einer kollektiven Mythe geworden.

**Tabelle 9: Fehlende Zusammenhänge**

Kindheit positiv	Geschlecht
Kindheit negativ	Aufenthalt, Geschlecht, Alter
Sozialisation	Geschlecht, Alter
Soziale Distanz	Geschlecht
Soziale Beziehungen	Geschlecht
Frauen als Kulturvermittler	Aufenthalt, Geschlecht, Alter
Negative Erfahrungen	Aufenthalt, Geschlecht
Anekdoten	Aufenthalt, Geschlecht, Alter
Indische Identität	Geschlecht
Niederländische Identität	Aufenthalt, Geschlecht, Alter
Ambivalente Identität	Aufenthalt, Geschlecht, Alter
Eigenschaften Indische Nederlanders	Aufenthalt, Alter
Eigenschaften Niederländer	Geschlecht, Alter
Transmission	Aufenthalt, Geschlecht, Alter
Reise negativ	Aufenthalt, Geschlecht, Alter
Reise positiv	Aufenthalt, Geschlecht, Alter
Bezug zur Vergangenheit	Aufenthalt, Geschlecht, Alter
Kein Bezug zur Vergangenheit	Aufenthalt, Geschlecht, Alter
Positiver Bezug zur Vergangenheit	Aufenthalt, Geschlecht, Alter
Negativer Bezug zur Vergangenheit	Aufenthalt, Alter
Historische Kontroversen	Aufenthalt, Geschlecht, Alter
Konfliktvermeidung	Aufenthalt, Geschlecht, Alter
Konflikt	Aufenthalt, Geschlecht, Alter
Umgebung: Niederlande	Aufenthalt, Geschlecht
Definition	Aufenthalt, Geschlecht, Alter

Bei 17 Paaren ergaben sich (siehe Tabelle 10) - mehr oder weniger stark ausgeprägte - Zusammenhänge<sup>31</sup>.

Dort zeigt sich zunächst, daß die Gruppe "Kindheit positiv" stark an einen persönlichen Aufenthalt in Indonesien gekoppelt ist. Dies gilt in - allerdings etwas schwächerem Maß - für "Alter", wobei dieser Zusammenhang allerdings als trivial zu gelten hat, da "Aufenthalt" und "Alter" aufgrund der starken Reduktion wiederum zusammenhängen ( $\Phi = .77^{**}$ ;  $\Lambda$

31. Als Zusammenhang wurden nur  $\chi^2$ -Werte mit  $p \geq .05$  akzeptiert.

**Tabelle 10: Themengruppen u. externe Variablen**

Gruppe	Alter	Geschlecht	Aufenthalt
Kindheit positiv	Phi $-.37^{**}$ lbd <sup>a</sup> .26	N/A	Phi $-.41^{**}$ lbd $-.35$
Sozialisation	N/A	N/A	Phi $-.41^{**}$ lbd .30
Soziale Distanz	Phi $-.45^{**}$ lbd .26	N/A	Phi $-.48^{**}$ lbd .35
Soziale Beziehungen	Phi $-.52^{**}$ lbd .31	N/A	Phi $-.55^{**}$ lbd .40
Negative Erfahrungen	Phi $-.53^{**}$ lbd .47	N/A	N/A
Indische Identität	ZNM <sup>b</sup>	N/A	ZNM
Eigenschaften INL	N/A	ZNM	N/A
Eigenschaften NL	N/A	N/A	Phi $-.33^{**}$ lbd .25
Negativer Bezug zur Vergangenheit	N/A	Phi $.46^{**}$ lbd .21	N/A
Umgebung: Niederlande	Phi $-.43^{**}$ lbd .37	N/A	N/A
Umgebung: Indonesien	Phi $-.43^{**}$ lbd .37	Phi $.33^{**}$ lbd .26	Phi $-.48^{**}$ lbd .45

- a. Lambda bezieht sich nur auf Alter, Geschlecht und Aufenthalt als abhängiger Variable.
- b. ZNM bedeutet, daß ein Zusammenhang aufgrund ungleicher Zellenverteilung (50% <5) nicht gemessen werden konnte.

(Aufenthalt *dep.*) = .75; (Alter *dep.*) = .74). Obwohl dieses Thema also immer wieder in den Gesprächen vorkommt, wird es nicht von den Jüngeren aufgenommen. Wie schon aus der qualitativen Analyse ersichtlich wurde, sind es ja gerade auch diese, die ein eher kritisches Verhältnis zu diesen Darstellungen der Älteren entwickeln.

Ebenfalls an das persönliche Erleben gekoppelt ist die Gruppe Sozialisation. Ich hätte hier einen stärkeren Zusammenhang mit "Alter" erwartet. Aber das Thema "Niederländisch zentrierte Erziehung" innerhalb der Gruppe hat hier wahrscheinlich den Ausschlag gegeben. Während dies von den Jüngeren in den Interviews erwähnt wurde, tauchte bei ihnen das Thema "Schulunterricht auf hohem Niveau" nie auf. Insgesamt werden in der Gruppe eher positive, an Leistung und Erfolg assoziierte Merkmale gebunden. Da diese aber nicht an Geschichten oder Anekdoten gekoppelt sind, sondern eher als abstrakte Ideen in den Kreisen der Älteren zirkulieren, wird die Weitergabe an die Jüngeren verhindert.

Zusammenhänge mit Aufenthalt und Alter ergeben sich auch mit der Gruppe "Soziale Distanz" (siehe Tabelle 11):

Tabelle 11:

SODISREC		Soziale Distanz by		AUFENTHA		Aufenthalt in Indonesien	
		AUFENTHA		Page 1 of 1			
Count						Row	
Col Pct		Ja		Nein			
						Total	
		1		2			
SODISREC							
	,00	12	19			31	
Kein Thema		52,2	95,0			72,1	
	1,00	11	1			12	
Ein oder mehrere Themen		47,8	5,0			27,9	
	Column	23	20			43	
	Total	53,5	46,5			100,0	
	Chi-Square		Value		DF		Significance
	Pearson		9,75214		1		,00179
	Minimum Expected Frequency -		5,581				
	Statistic		Value		ASB1	Val/ASB0	Approximate Significance
	Phi		-,47623				,00179 *1
	Lambda :						
	symmetric		,21875		,15409	1,28100	
	with SODISREC dependent		,00000		,00000		
	with AUFENTHA dependent		,35000		,22444	1,28100	
	*1 Pearson chi-square probability						
	Number of Missing Observations: 0						

Rund die Hälfte der Informanten (46,5%) kann keine persönlichen Erlebnisse in Indonesien vorweisen. Nur einer von ihnen (5%) erwähnt ein Thema aus dieser Gruppe. Auf der anderen Seite machen 47,8 % (n=11)

der Informanten, die einen längeren Aufenthalt in Indonesien erlebt haben, Themen aus der Gruppe zum Bestandteil ihrer Ausführungen. Ähnlich ist die Verteilung im Hinblick auf das Alter: Zwar findet sich das Thema elfmal bei Informanten der älteren und mittleren Generation, aber nur einmal bei einem der Jüngeren. Eine mögliche Erklärung könnte sein, daß die Weitergabe in diesem Fall auf zwei Wegen erfolgen kann: Zum einen mit Hilfe von Anekdoten, die, durchaus an die Jüngeren weitergegeben werden. Als diskrete Aussage formuliert, so wie es die Kategorien- definition vorsah, bleibt es ein Thema der Älteren bzw. derjenigen, die persönliche Erfahrungen in Indonesien gemacht haben.

Ein Ergebnis der qualitativen Analyse kann man aufgrund der quantitativen nicht bestätigen: Die Themengruppe "Soziale Beziehungen" hängt nicht mit dem Geschlecht zusammen. Statistisch gesehen, läßt sich dies nur für "Aufenthalt" und "Alter" beweisen (siehe Tabelle 12):

Tabelle 12:

SOZBEZRE Soziale Beziehungen		by AUFENTHA Aufenthalt in Indonesien	
		Page 1 of 1	
Count		AUFENTHA	
Col	Pct	Ja	Nein
		1	2
		Row Total	
SOZBEZRE			
	,00	12	20
Kein Thema		52,2	100,0
	1,00	11	
Ein oder mehrere Themen		47,8	
		23	20
Column Total		53,5	46,5
			100,0
Chi-Square		Value	DF
Pearson		12,85326	1
Minimum Expected Frequency -		5,116	
Statistic		Value	ASE1
Phi		-,54673	
Lambda :			Val/ASE0
symmetric		,25806	,15718
with SOZBEZRE dependent		,00000	,00000
with AUFENTHA dependent		,40000	,21909
Approximate Significance			
			,00034 *1
			1,44830
			1,44830

Diesmal sind die Zahlen noch deutlicher als in Tabelle 11. Haben die Befragten keine persönlichen Erfahrungen in Indonesien gemacht, kommt das Thema in ihren Ausführungen nicht vor (100%). Bei "Alter" ist der Zusammenhang wieder etwas schwächer, aber immer noch hinreichend stark. Weniger als das Geschlecht ist also die persönliche Erfahrung entscheidend dafür, wie sehr das Thema "Soziale Beziehungen" Be-

standteil der biographischen Darstellung wird. Eine mögliche Erklärung ist vielleicht, daß die Erfahrung von Not und Schwierigkeiten im Heimatland die Aktivierung familiärer Netze besonders dringlich machte, und daß deswegen die Rekonstruktion von Beziehungen, so wie sie in Kategorie 1.5.2. ausgedrückt wurde, von den Frauen als Botschaft weitergegeben wird. Allerdings ist ein Zusammenhang zwischen den Gruppen "Negative Erfahrungen" und "Soziale Beziehungen" nicht nachweisbar. Zudem muß berücksichtigt werden, daß durch die starke Datenreduktion, die durch die Gruppierung der Themen entstanden ist, sich die Grenzen zwischen männlichen und weiblichen Themen ("Geschichte der Ahnen" und "Genealogische Rekonstruktion") verwischt haben.

Eindeutig bestätigt werden kann allerdings der in der qualitativen Analyse ausgemachte Zusammenhang zwischen "Negativen Erfahrungen" und "Alter" (siehe Tabelle 13):

Tabelle 13:

NEGREC Negative Erfahrungen by GEN2 Alter		GEN2		Page 1 of 1	
Count		Altere u.		Junge	
Col	Pct	Mittlere			Row
		1,00	2,00		Total
NEGREC					
	,00	3	12		15
Kein Thema		12,5	63,2		34,9
	1,00	21	7		28
Ein oder mehrere Themen		87,5	36,8		65,1
		24	19		43
	Total	55,8	44,2		100,0
	Chi-Square		Value	DF	Significance
	Pearson		11,98059	1	,00054
	Minimum Expected Frequency -		6,628		
	Statistic		Value	ASE1	Val/ASE0
	Phi		-,52784		,00054 *1
	Lambda :				
	symmetric		,41176	,17779	1,91508
	with NEGREC dependent		,33333	,23727	1,16504
	with GEN2 dependent		,47368	,14788	2,48506

\*1 Pearson chi-square probability  
Number of Missing Observations: 0

Zwar hat das Sample hier ein kleines Ungleichgewicht, weil mit 55,8% (n=24) die Mehrheit der Informanten in die Gruppe der Älteren und Mittleren fällt, und 87,5% (n=21) der Informanten aus dieser Gruppe Themen aus dem Bereich der "Negativen Erfahrungen" erwähnen. Im Gegensatz dazu die Jüngeren: Nur 7 (36,8%) von ihnen wissen hier etwas dazu zu sagen. Aus der qualitativen Analyse weiß man jedoch, daß dies oft

nur Bruchstücke, also Informationsfetzen sind, die keineswegs in Dringlichkeit und Qualität der Ausführung die Darstellungen der Älteren erreicht.

Beim Thema "Indische Identität" ist der Zusammenhang zwischen "Aufenthalt" einerseits und "Alter" auf der anderen Seite zwar aufzeigbar, aber äußerst schwach. Aufgrund der Tatsache, daß in der Tabelle zwei von vier Feldern mit einer erwarteten Häufigkeit von unter 5 besetzt sind, können keine zuverlässigen Zahlen ermittelt werden. Wenn das Thema jedoch erwähnt wird, dann nie von den Jüngsten, und in keinem Fall von denen, die keine persönliche längere Erfahrung in Indonesien gesammelt haben. Auch dies ist schon als Ergebnis der qualitativen Analyse festgestellt worden.

Ein ähnlich schwacher Zusammenhang ist hinsichtlich der Gruppe "Eigenschaften Indischer Niederländer" und "Geschlecht" wahrzunehmen. Auch hier muß allerdings auf solide Maße verzichtet werden. Diese Verteilung ermutigt allenfalls zur Aufstellung einer Hypothese für zukünftige Untersuchungen. Die Tatsache, daß in diesem Fall keine Zusammenhänge zwischen "Alter" und "Aufenthalt" festgestellt werden konnten, bestätigt ebenfalls die entsprechende Feststellung der qualitativen Analyse.

Die Gruppe "Eigenschaften der Niederländer" hängt nur schwach mit der Variable "Aufenthalt" zusammen. Eine Erklärung hierfür könnte sein, daß erst persönliche Erfahrung dieses Thema zum Bestandteil der individuellen Mythe werden läßt. Das Erzählen als Mittel der Weitergabe scheint nicht auszureichen.

Ein Zusammenhang besteht zwischen der Gruppe "Negativer Bezug zur Vergangenheit" und "Geschlecht" (siehe Tabelle 14).

Einerseits gilt zwar auch hier, daß die schwache Besetzung der unterschiedlichen Zellen zu äußerster Vorsicht bei der Interpretation der Zahlen Anlaß gibt, andererseits taucht das Thema in keinem Fall bei einem Mann auf. 37,5% (n=9) der Frauen aber haben einen negativen Bezug zur Vergangenheit. Da es jedoch keine Zusammenhänge zwischen "Alter" und "Geschlecht" sowie "Aufenthalt" und "Geschlecht" gibt - das Sample ist hier homogen - fällt eine Erklärung schwer. Vielleicht ist für die Frauen der Wechsel von der Heimat in die neue Umgebung mit besonderen Schwierigkeiten verbunden gewesen. Schwierigkeiten, die nicht unbedingt in den Gesprächen thematisiert wurden, wie Statuswechsel oder -verlust, Ereignisse in ihrem Leben, die in besonderem Maße zur Reflexion anregen.

Die überwiegend schlechten Erfahrungen, die in der Themengruppe "Umgebung: Niederlande" abgefangen werden - allesamt Bereiche, die

Tabelle 14:

VERGAREK	Negativer Count Col Pct	Bezug zur SEX male	Vergangenheit female	by SEX Row Total	Geschlecht	
VERGAREK		1	2	Total		
Kein Thema	,00	19	15	34		
		100,0	62,5	79,1		
Ein oder mehrere Themen	1,00		9	9		
			37,5	20,9		
	Column Total	19	24	43		
		44,2	55,8	100,0		
	Chi-Square		Value	DF	Significance	
Pearson			9,01103	1	,00268	
Fisher's Exact Test:						
One-Tail					,00232	
Two-Tail					,00248	
Minimum Expected Frequency -			3,977			
Cells with Expected Frequency < 5 -			1 OF	4 ( 25,0%)		
	Statistic		Value	ASE1	Vai/ASE0	Approximate Significance
Phi			,45778			,00268 *1
Lambda :						
symmetric			,14286	,19280	,68978	
with VERGAREK dependent			,00000	,00000		
with SEX dependent			,21053	,27268	,68978	
*1 Pearson chi-square probability						
Number of Missing Observations: 0						

sich mit den Schwierigkeiten nach der Wanderung auseinandersetzen, stehen in einem Zusammenhang mit "Alter". Obwohl dieses Thema also von nahezu jedem "Älteren" angesprochen wird, wird es nur bei weniger als der Hälfte der "Jungen" aufgegriffen, bzw. zum Bestandteil der Darstellung.

Warum schließlich die Frauen der Vergangenheit kritischer gegenüber stehen, könnte sich aus den Themen der Gruppe "Umgebung: Indonesien" und den Zusammenhängen mit "Alter", "Geschlecht" und "Aufenthalt" ergeben.

Der stärkste Zusammenhang ergibt sich hier mit der Variable "Aufenthalt" (siehe Tabelle 15). Die Themengruppe ist überwiegend mit positiven Kategorien besetzt, wie dem luxuriösen Leben, dem materiellen Wohlstand. Besonders eindringlich scheint dieses auch den Frauen gewesen zu sein. Immerhin wird bei 75% von ihnen diese Themengruppe zur Darstellung der Familiengeschichte herangezogen. Die Erklärung ergibt sich aus dem ethnographischen Kontext: Die in Indonesien lebenden Frauen hatten

Tabelle 15:

INDOREC2 Umgebung: Indonesien by AUPENTHA Aufenthalt in Indonesien		AUPENTHA		Page 1 of 1	
INDOREC2	Count	AUPENTHA		Row	
	Col Pct	Ja	Nein	1	2
				Total	
		1	2		
INDOREC2					
	,00	4	13		17
Kein Thema		17,4	65,0		39,5
	1,00	19	7		26
Ein oder mehrere Themen		82,6	35,0		60,5
	Column	23	20		43
	Total	53,5	46,5		100,0
	Chi-Square		Value	DF	Significance
	Pearson		10,14324	1	,00145
	Minimum Expected Frequency -		7,907		
	Statistic		Value	ASE1	Val/ASE0
	Phi		-,48568		,00145 *1
	Lambda :				
	symmetric		,40541	,17117	1,97356
	with INDOREC2 dependent		,35294	,21161	1,37064
	with AUPENTHA dependent		,45000	,15289	2,31484
	*1 Pearson chi-square probability				
	Number of Missing Observations: 0				

gegenüber denen, die in Europa lebten einen enormen Vorteil: Befreit von einer Belastung durch Pflichten im Haushalt, konnten sie sich Beruf oder uneingeschränkt der Verwirklichung ihrer persönlichen Interessen widmen. Daß dies vor allem eine Frage des Alters ist, zeigt die hohe Ausprägung von Phi in Tabelle 10.

## Multivariate Verfahren

### Clusteranalyse

Mit Hilfe der Clusteranalyse werden Gruppen von Personen oder Objekten gebildet (Backhaus 1987: 115), die eine hohe Ähnlichkeit miteinander aufweisen. Zu diesem Zweck wird die Datenmatrix "Informanten-Variablen" dichotomisiert. Das bedeutet, daß in dieser Matrix nur noch abgebildet wird, ob ein bestimmtes Thema bei einem Informanten vorkommt oder nicht. Für die anschließenden Verfahren wurden wieder alle 88 möglichen Kategorien berücksichtigt. Anschließend wurde diese Matrix in zwei Ähnlichkeitsmatrizen umgewandelt, jeweils eine für die Informanten, die andere für die Variablen. Dabei wurde die Ähnlichkeit über ANTHROPACs (Borgatti 1992) Prozedur *similarities-positive*

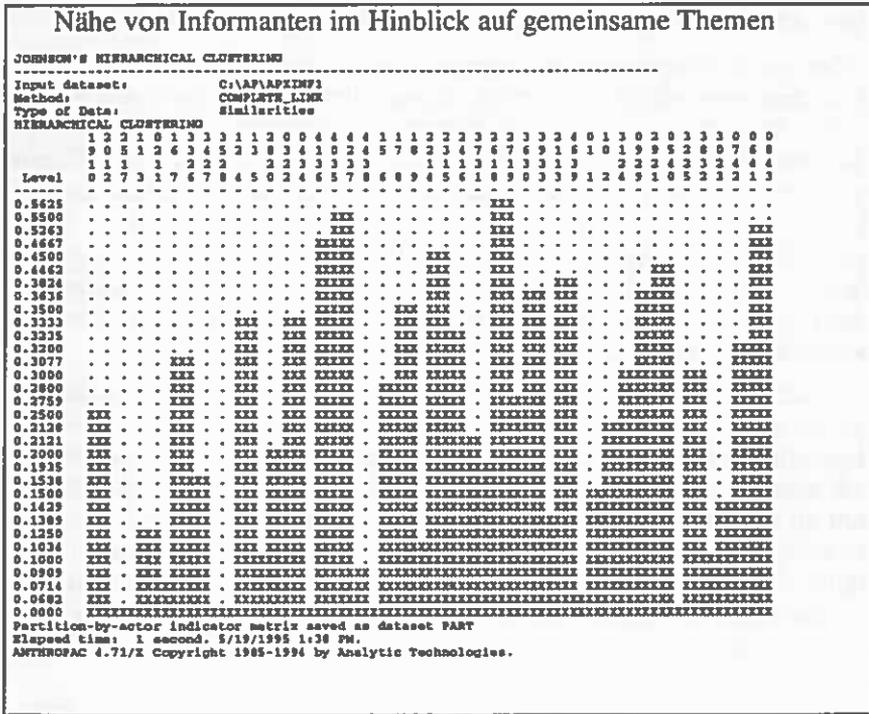
*matches* ausgedrückt. Bei diesem Verfahren, das in der Literatur auch Tanimoto-Koeffizient genannt wird (Backhaus 1987: 120), wird die Ähnlichkeit eines jeden Paares von Informanten in dem einen, und Kategorien im anderen Fall anhand der Überprüfung gemeinsamer Eigenschaften im Hinblick auf das *Auftreten* eines Merkmals gemessen. Dabei wird die Summe gemeinsamer Eigenschaften durch die Summe von gemeinsamen Eigenschaften und den individuellen Eigenschaften jeder Paarhälfte geteilt. Somit bleibt allerdings die Ähnlichkeit eines Paares, die sich aus *nicht vorhandenen* Eigenschaften ergibt, unberücksichtigt. Das sich anschließende hierarchische Clusterverfahren nach Johnson bietet mehrere Algorithmen an, um die Nähe der Personen oder Objekte abzubilden. Im folgenden wurde das *Complete-Linkage-Verfahren* eingesetzt (Backhaus 1987: 140). Es läßt sich verallgemeinernd sagen, daß im Gegensatz zum *Single-Linkage-Verfahren* diese Prozedur zur Bildung kleinerer Gruppen führt. Ich habe beide Verfahren getestet, und letzteres schien mir für die weitere Analyse nicht ergiebig zu sein.

Nach diesem Verfahren können nun mehrere Gruppen gebildet werden (vergleiche Abbildung 11):

- A: Meike (19; 18)<sup>32</sup>, Richie (20; 19)
  - B: Ellie (16; 15), Asta (33; 32), Lilian (34; 33), Jane (11; 11), Eveline (02; 02)
  - C: Woody (12; 12), Laila (13; 13), George (03; 03), Henry (04; 04), Toni (28; 27)
  - D: Anna (40; 39), Henriëtte (42; 41), Erna (41; 40), Rein (44; 42)
  - E: Antje (26; 25), Ruud (27; 26), Pia (22; 21), Oscar (23; 22), Gabriella (24; 23), Edmond (37; 36), Sjakkie (21; 20), Cleo (46; 43), Astrid (36; 35), Olga (39; 38), Harko (17; 16), Margriete (18; 17), Isabella (15; 14)
  - F: Saskia (29; 28), Marlene (05; 05), Jaqueline (09; 09), Simone (31; 30), Michel (10; 10), Klaas (01; 01)
  - G: Ivo (06; 06), Gerard (08; 08), Kurt (07; 07), Emiel (30; 29), Didier (32; 31), Tania (38; 37)
- Isoliert: Andrew (25; 24) und Gaston (35; 34)

32. In Klammern steht an erster Stelle das Informantenetikett, das sich in den folgenden Clusteranalysen in der oberen Reihe wiederfindet. Dahinter, fett gesetzt, die Fallzahlen, mit denen die Informanten in den Abbildungen der univariaten Analysen dargestellt werden. Die Abweichungen ergeben sich aus der Tatsache, daß bei der Auswertung der Interviews einige Informanten zwar aufgrund ihrer Anwesenheit einen Kode erhielten, dann aber als "Schweiger" (INF14), oder sonstigen Gründen ausgeschlossen wurden (INF43 der Regisseur des Videobandes; bei INF45 war das Band aus technischen Gründen unbrauchbar).

Abbildung 11



Zunächst einmal ist auffällig, daß es ganz offensichtlich keinen Zusammenhang zwischen diesen Gruppen und den Besetzungen der Interviews gibt. Inhaltlich stehen sich in den meisten Fällen ganz andere Informanten nahe als durch die Gesprächssituation. Dies ist meines Erachtens ein weiterer Indikator für die Produktivität der Gruppendiskussion, die eben nahezu keine Konsenseffekte erzeugt hat. Die Isolation von Andrew und Gaston ist trivial und hängt mit dem geringen Redeanteil zusammen. Was verbindet nun die Mitglieder der einzelnen Gruppen nach dem bisherigen Stand der Analyse?

Gruppe A wird durch ein Geschwisterpaar gebildet, das zwar zur Familie gerechnet wird, aber bezüglich seines ethnischen Hintergrunds mit den anderen wenig gemeinsam hat. Die beiden sind in Surinam aufgewachsen und von diesem Umstand viel mehr geprägt als ihre in den Niederlanden aufgewachsenen Altersgenossen. So gehen sie im Interview ausführlich auf ihre Kindheit und Jugend in diesem Land ein und beziehen

die Fragen des Interviewers mehr auf diese Situation. Ihre Anbindung über die in Indonesien aufgewachsene Mutter und die erweiterte Familie an einen *indischen* Kontext spielt in ihren Ausführungen keine Rolle.

Gruppe B könnte man als Sympathisanten ohne direkten Bezug zur ethnischen Anbindung bezeichnen. Keiner dieser Informantinnen nimmt explizit für sich in Anspruch, kompetent im Hinblick auf die ethnische Situation oder die Familiengeschichte zu sein. Sie stehen diesem Thema aber auch nicht kritisch oder ablehnend gegenüber. Ein Blick auf die Abbildungen der univariaten Analysen zeigt, daß sich die hohe Nähe zwischen Ellie und Asta hauptsächlich aus ihrem gemeinsamen Wissen um *Eigenschaften Indischer Niederlanders* ergibt. Jane und Eveline stoßen erst relativ spät im Diagramm zu den anderen - das verbindende Element zwischen ihnen scheint die Schweigsamkeit zu sein.

Anders präsentiert sich Gruppe C, die man als "die Unschlüssigen" bezeichnen könnte. Diese Informanten reflektieren über ihren ethnischen Hintergrund, ohne allerdings zu einer Entscheidung gekommen zu sein. Sie sind, was die Geschichte der Familie angeht, kompetent und interessiert, mehr zu erfahren. Hinzu kommt ein eher formaler Wesenszug: Es handelt sich ausnahmslos um Informanten mit sehr geringen Redebeiträgen. Man könnte also auch hier einen Intervieweffekt entdecken, der abhängig von der gelieferten Textmenge ist. Schließlich stehen sich zwei Geschwisterpaare besonders nahe: Es handelt sich um Woody und Laila sowie Henry und George.

Wie bei Gruppe C könnte auch bei den Informanten der Gruppe D ein Intervieweffekt aufgetreten sein. Es sind die Akteure des Videobandes. Dieses Gespräch wurde durch Marlene extrem gelenkt, so daß sich alle vier zu bestimmten, vorgegebenen Themen äußerten. Außerdem wurde das Band durch Schnitte nachträglich manipuliert. Die Clusteranalyse macht diesen Effekt überzeugend deutlich. Aber auch vor dem ethnographischen Hintergrund macht die Konfiguration einen Sinn. In diesen vier Alten vereint sich der gesamte Erfahrungsschatz der untersuchten Familie. Sie sind es, die den größten Teil ihres Lebens in der ursprünglichen Heimat zugebracht haben und alle Höhen und Tiefen bewußt miterlebt haben. Es scheint also so betrachtet, auch nicht unlogisch, daß sich in den Darstellungen dieser vier Menschen gleichsam die Essenz einer gemeinsamen Mythe widerspiegelt.

Die größte Gruppe, E, wird von Informanten gebildet, die sich intensiv mit der Geschichte der Familie, und mit dem eigenen Hintergrund auseinandergesetzt haben. Sie alle eint auch die kritische Distanz zur Vergangenheit, dem kolonialen Milieu und der erfolgten Wanderung. Für diese

Menschen ist die Vergangenheit abgeschlossen. Das gilt in sicherem Maße für Antje und Ruud, aber auch Harko und Margriete, bei denen aber aufgrund des über weite Strecken im intensiven Dialog stattgefundenen Gespräches ein hoher Konsens über eine Betrachtung der eigenen Vergangenheit eingetreten ist. Pia und Oscar dagegen verbindet das Wissen um die *Eigenschaften der Indischen Nederlanders*, die *Anekdoten*, die Gruppe *Umgebung: Niederlande*, und vor allem die *Negativen Erfahrungen*, die sie nahezu gleich erlebten. Bei Sjakkie und ihrer Mutter Cleo werden mit Sicherheit eine Reihe von gleichen Themen angesprochen, weil die beiden eine gemeinsame Interviewsituation erlebten, andererseits liegt ein Gespräch zwischen Sjakkie und Marlene vor, so daß es nicht plausibel erscheint, daß das hohe Maß an Nähe einzig auf diesen Effekt zurückzuführen ist. Neben anderen Übereinstimmungen erweisen sich Mutter und Tochter als Experten für *Soziale Beziehungen*. Astrid und Olga treffen sich in der gleichen Einschätzung der Umstände in der *Umgebung: Indonesien* und der *Umgebung: Niederlande*.

Die Informanten der Gruppe F sind im höchsten Maße kompetent, und auch sie haben sich mit dem eigenen Hintergrund und der Vergangenheit intensiv auseinandergesetzt. Aber ihre Konsequenz ist eine andere. Sie haben ihre Andersartigkeit als Bestandteil ihrer Biographie und ihrer Auffassung vom Leben, positiv akzeptiert und zum Teil in beträchtlichem Ausmaß in ihr Leben integriert. Das gilt besonders für Marlene und Saskia, deren Nähe sich aus einer Vielzahl von gemeinsamen Themen ergibt. Das sind vor allem die Kenntnisse über *Anekdoten*, aber auch die *negativen Erfahrungen*. Ein Rückblick in die qualitative Analyse zeigt darüber hinaus, daß es das weitestgehend gleiche Einstellungssystem ist, wie es sich in der Gruppe *Eigenschaften Indischer Nederlanders* ausdrückt, das diese Frauen verbindet.

Die letzte Gruppe, G, ist von Informanten besetzt, die man als "stille Beobachter" bezeichnen könnte. Sie grenzen sich einerseits deutlich von den anderen ab und sind nicht Teil der *indischen "Gemeinschaft"* geworden. Ihre Haltung ist neutral bis distanziert, jedoch nie ohne Sympathie. Dies gilt für Gerard und seinen Sohn aus erster Ehe, Ivo, die in der Gruppe am engsten verbunden sind. Kurt, Marlenes und Gerards Sohn dagegen, stößt erst später dazu. Er kann, wie aufgezeigt wurde, weit mehr mit den Ideen seiner Mutter anfangen, als dies den anderen beiden möglich ist. Hinzu kommt seine persönliche Unsicherheit. Die Nähe dieser drei ergibt sich aus dem gemeinsamen Wissen über *Eigenschaften Indischer Nederlanders*.

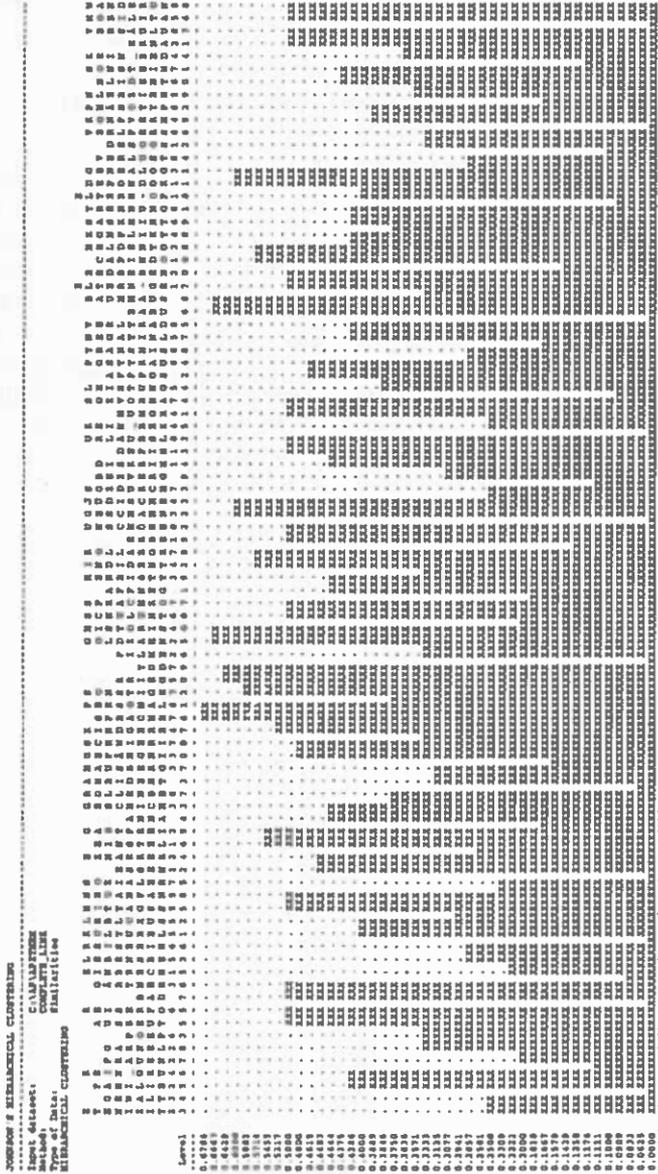
Nach dem gleichen Verfahren kann nun auch die Nähe der Themen analysiert werden (vergleiche Abbildung 12):

Wieder wurden Gruppen gebildet:

- A: Ethnizität, Gewalt
- B: Familie, Kein Grund zur Idealisierung, Frauen als Kulturvermittler, Garnelenjagd, Poep, Ausbeutung, Reise erfolglos
- C: Bad, Gärtner, Einbrecher, Lebensbedingungen ideal, Reiseeinstellung: offen und unbelastet
- D: Kulturelle Aktivitäten, Leben im Luxus, Mittagsschlaf, Neuanfang, Sozialer Aufstieg
- E: Essen, Externe Marker, Einstellung, Gastfreiheit, Anna, Geschichten, Rollenbesetzung
- F: Ambivalente Identität, Haus in den Bergen, Gefängnis, Schwierigkeiten - kulturelle/klimatische Ursachen, Kindgerechte Aktivitäten, Sprache, Personal, Sonstiges, Krieg, Tod
- G: Filme, Goldzahn, historische Kontroversen, Schulunterricht auf hohem Niveau, Sprachliche Kompetenz, Affekte - negativ, Henriëtte, Indische Identität, Kollaboration, Erbe, Unschöne Kindheit.
- H: Geschichte der Ahnen, Jüdische Vorfahren, Soziale Distanz durchbrochen, Bezug zur Vergangenheit, Diskriminierung, Definition, Urlaubsland
- I: Kein Bezug zur Vergangenheit, Humor, Soziale Distanz vorhanden, Leistung, Affekte positiv, Posttraumatisches Schocksyndrom, Transmission, Negative Einstellung der Niederländer, Verlust der Tradition
- J: Babu, Baumhaus, Elternbild - Geschlossenes Einstellungssystem, Erdbeben, Cap Indo, Heldentod, Konflikt, Der Schuß in Saskias Hintern, Trennung, Elternbild: offenes Einstellungssystem, De Pondok, Genealogische Rekonstruktion, Verlust, De Soos, Verlust des Personals
- K: Konfliktvermeidung, Phänotyp, Niederländisch-zentrierte Erziehung, Niederländische Identität, Sonstige Einstellungen, Kein Bedarf, Erna, Versäulung
- L: Mobilität, Wanderung

Die durch die Clusteranalyse dargestellte "Nähe" der Themen sagt nichts über einen kausalen oder inhaltlichen Zusammenhang aus. Aber schon die Nähe, die sich über die Nennung durch die Informanten ergibt, überrascht bei einigen Beispielen (im folgenden wurden einige sehr dicht stehende

Abbildung 12 Nähe von Themen



Paare, auf den "Spitzen" der Türme in der grafischen Darstellung betrachtet):

Die geringste Entfernung findet sich zwischen den Themen *Sprache* und *Personal*, also einerseits einer als "typisch" empfundenen Eigenschaft Indischer Nederlanders, andererseits einem prägenden Moment der kolonialen Umgebung. Dadurch wird die Wichtigkeit dieser beiden Themen im Bewußtsein der Informanten für die Beschreibung ihrer Situation ausgedrückt. Wenn also ein Informant das Thema *Sprache* erwähnt, dann taucht mit großer Sicherheit auch das Thema *Personal* in seinen Ausführungen auf.

Auffällig sind weiterhin Assoziationen zwischen Personen und bestimmten Begriffen: Hier findet sich *Henriëtte-Affekte negativ*, *Erna-Versäulung* und *Anna-Geschichten*, das heißt, an jede der Personen wird ein bestimmtes Thema gekoppelt. Überraschend sind Verbindungen wie zwischen *Indische Identität-Kollaboration* und *Ausbeutung-Reiseeinstellung offen*. Es scheint, daß ein jeweils negatives, die Vergangenheit kritisch reflektierendes Moment vorhanden sein muß, um Momente der Offenheit und Identifikation zuzulassen. Auf der anderen Seite wird dann auch das Paar *Niederländische Identität-Sonstige Einstellungen* analog gebildet (letzteres fing eher negative Eigenschaften der Niederländer ab). Andere Paare überraschen dagegen nicht. So *Kulturelle Aktivitäten-Leben im Luxus*, *Krieg-Tod*, *Einstellung-Gastfreiheit*, *Geschichte der Ahnen-Jüdische Vorfahren*, *Negative Einstellungen der Niederländer-Verlust der Tradition*, *Cap Indo-Heldentod* und *De Pondok-Genealogische Rekonstruktion*.

Ich meine, in den Gruppen folgende Gemeinsamkeiten zu entdecken:

- A) Diese Gruppe zeichnet sich durch hohe *Konfliktladung* aus. Einerseits umfaßt sie Themen, in denen Informanten bei anderen eine Steigerung ethnischen Bewußtseins ausmachen. Hinzu kommt das Moment erlebter *Gewalt* zwischen Niederländern und Indischen Nederlanders nach der Wanderung. Beide Momente sind typische Anti-Werte indischer Kultur. Die qualitative Analyse zeigte, daß für viele Informanten die *Anpassungsfähigkeit* ein wichtiges Merkmal der Gruppe ist. Ausgesprochene Scheu vor Konflikten wurde oftmals als typisch indische Eigenschaft beschrieben. So scheint die Berufung auf eine indische Identität, die dann auch noch offensiv nach außen getragen wird, für die Befragten Anlaß zu sein, bei anderen Passagen des Gespräches indirekt auf die ihrer Meinung nach möglichen Konsequenzen eines solchen Verhaltens zu verweisen.
- B) Bei diesem Cluster scheint es sich um eine Gruppe weiblicher Themen

- zu handeln. Diese haben entweder Frauen zum Inhalt oder werden von Frauen vorgebracht. So werden Themen wie *Familie* und *Frauen als Kulturvermittler*, aber auch die skeptischen Momente, in denen eine Weigerung zum Ausdruck kommt, die Vergangenheit in allzu rosigem Licht zu sehen, in einen Zusammenhang gebracht. Je kritischer dies gesehen wird, etwa durch das Thema *Ausbeutung*, desto häufiger wird dann auch vom Befragten die Möglichkeit durch eine Reise ins Heimatland seinen Wurzeln auf den Grund zu gehen, abgelehnt.
- C) *Unkritische Reflexion* ist das gemeinsame Merkmal dieser Gruppe. Fröhliche Anekdoten, zwei davon sind Kinderstreiche, werden durch den Eindruck von für *Kindern idealen Lebensumständen* ergänzt. Daran eng geknüpft ist eine von der Vergangenheit und der ethnischen Anbindung unbelastete Reiseeinstellung, die nicht die Klärung des ethnischen Standpunktes zum Inhalt hat.
- D) Im Mittelpunkt der nächsten Gruppe stehen Ausführungen, in denen *sozialer Status* verhandelt wird, oder die sich auf gesellschaftliche Positionen und Stellungen beziehen. Zugegebenermaßen läßt sich die Einordnung der Anekdote um den *Mittagsschlaf* hier nicht sinnvoll interpretieren. Die anderen Themen stehen inhaltlich jedoch in einem engen Zusammenhang. Werden also Themen angesprochen, mit denen Informanten ihre hohe Position, oder ihren materiellen Wohlstand in Indonesien beschreiben, so sind in einer Vielzahl von Fällen damit Schilderungen um Statusrekonstruktionen in den Niederlanden nach erfolgter Wanderung verbunden.
- E) *Indische Eigenschaften* bilden den Schwerpunkt einer weiteren Gruppe. Von 13 zur Verfügung stehenden Ausprägungen werden sechs in eine Beziehung gebracht. *Essen, externe Marker, Einstellung* und *Gastfreiheit* erwiesen sich auch in der qualitativen Analyse als die wichtigsten gemeinsamen Nenner in den Schilderungen der Befragten. Sie werden mittelbar mit Beschreibungen der Person Annas gekoppelt, die man als eine der zentralen Figuren aus dem Korpus kennengelernt hat.
- F) Schilderungen *indischen Alltags*, die sich chronologisch in den Zeitraum der japanischen Besetzung, des *bersiaps*, und des Übergangs nach der Wanderung einordnen lassen, machen den Charakter dieser Gruppe aus. Zum einen werden die *negativen*, persönlich belastenden, *Erfahrungen* dargestellt. Dann finden sich aber auch positiv besetzte Erinnerungen, wie das für alle Befragten der mittleren und älteren Generation so wichtige *Haus in den Bergen*, jenem Hort geselligen Zusammenseins, der die Beteiligten mindestens genauso zusammenschweißte wie die Schrecken der japanischen Lager.

- G) Das verbindende Element dieser Gruppe stellt unter Umständen ein *Intervieweffekt* dar. Es sind Themen, wie sie in der Verbindung untereinander einzigartig im Gespräch zwischen Marlene und Simone vorkamen. Simone, das zeigte sich in der qualitativen Analyse, nimmt eine Außenseiterposition ein, in dem sie versucht, das von Marlene produzierte positive, in ihren Augen zu rosige Bild der Vergangenheit zu relativieren, und eine Gegenposition dazu aufbaut. Warum sie dies tut, ist einigermaßen unklar: Mit hoher Wahrscheinlichkeit will sie dem Interviewer gegenüber so etwas wie Objektivität der Darstellung konstruieren. Dem steht aber jene Passage gegenüber, in der sie (unter *Sonstiges* kodiert) den Verlust der Heimat bedauert und wiederum einen Teil ihrer eigenen Aussagen relativiert.
- H) Themen, die ein *kritisches Auseinandersetzen* der Informanten mit ihrer eigenen, oder der kollektiven *Vergangenheit* darstellen, finden sich in dieser Gruppe. Der Bogen spannt sich von den Anfängen der Familie in Indonesien, wo darüber reflektiert wird, wie die Vorfahren überhaupt in dieses Land kamen. Er spiegelt die eigene Lebenserfahrung in Indonesien wider, wo etwa unter *Soziale Distanz durchbrochen* über das Verhältnis zwischen kolonialer und einheimischer Bevölkerung nachgedacht wird und vereint aber auch die Erfahrungen, die nach der Migration in der neuen Heimat gesammelt wurden, wie Momente der *Diskriminierung*. Es erstaunt nicht, daß die Frage nach der *Definition*, ein ausgesprochen selbstreflektierendes Moment, hiermit eng verbunden ist. In diesem Kontext wird das Problem der Erreichbarkeit des Heimatlandes entspannt aufgelöst: Das Verhältnis zur Geschichte ist verarbeitet, und eine Reise deswegen kein Problem. Indonesien kann als *Urlaubsland* erfahren werden, die Heimat ist anderswo.
- I) Im Gegensatz dazu gruppieren sich Themen im nächsten Cluster: Für die Informanten, die sich mit ihrer Hilfe ausdrücken, existieren die alten Verhältnisse unverändert weiter. Sie können *keinen Bezug zur Vergangenheit* haben, da sie keinerlei selbstkritischen Abstand zur eigenen Geschichte und ihren Erfahrungen entwickelt haben. Die Erlebnisse in Indonesien sind für diese Menschen unverständlich und rätselhaft geblieben. Aus ihrer Sicht wurde das Land aufgebaut und entwickelt. Sie selbst erlitten im Krieg und im *bersiap* Schreckliches. An die niederländische Umgebung haben sie sich nicht gewöhnt und empfinden es außerdem als Belastung, daß die Erfahrungen, die sie gemacht haben, nicht an die Nachkommen weitergegeben werden können.
- J) *Anekdoten* hängen in der Darstellung der Informanten stark zusammen. Wird eine erwähnt, so ist es sehr wahrscheinlich, daß dieser Erinnerung

noch eine Reihe von weiteren Geschichten folgen. An die Schilderung von Ereignissen ist interessanterweise auch die Darstellung der verwandtschaftlichen Beziehungen geknüpft. Geschichten, so kann man vermuten, sind ohne die Schicksale der damit verbundenen Personen, und deren Verwandten und Familienangehörigen nicht rekonstruierbar, sie bilden Assoziationsanker bzw. Knotenpunkte in einem Netz sozialer Beziehungen.

- K) Themen des *Identitätsmanagements* gruppieren sich um Momente der Erziehung, der (niederländischen) Identität und Anpassungsstrategien. Fühlen Informanten, daß sie in der Heimat das nötige Rüstzeug für den Übergang in das neue Land bekamen, so sind sie in der Lage, mit einem quasi ethnologischen Blick die andere Gesellschaft zu erfassen. Sie haben sich eingeordnet und verspüren dann auch *keinen Bedarf* zurückzukehren oder das Land der Ahnen und der eigenen Vergangenheit auch nur zu bereisen. Ihr Brückenkopf, ihre Schnittstelle zur Vergangenheit, Symbol des Wandels und der Anpassung an die neuen Verhältnisse in Indonesien ist *Erna*, die sich repräsentativ für die Bedürfnisse der restlichen Familie eine neue Existenz in Bandung aufbaute.
- L) *Wanderung* als Schicksal und als Ausdruck einer selbstdefinierten Identität wird zwar von den Informanten unterschiedlich bewertet, aber in einen assoziativen Zusammenhang mit einem Gefühl des Stolzes und der kulturellen Überlegenheit und Flexibilität gegenüber dem Kontext gebracht. Nicht ohne Grund: Denn, so fühlen sie, die durchstandenen Leiden, die Erfahrungen, die mit schweren Verlusten verbunden sind, können letztlich in modernen Kontexten, wo Mobilität und räumliche Ungebundenheit als Gradmesser persönlicher Leistungsfähigkeit gelten, zu Tugenden werden.

### Vergleich der Abbildungsfähigkeit durch Worthäufigkeiten

Im folgenden stellt sich nun die Frage, inwieweit die Erkenntnisse aus dem Text und die Ergebnisse über die Zusammenhänge zwischen den Informanten auch auf anderem Wege als über die inhaltliche Verkodung des Textes gewonnen werden können. Zu diesem Zweck wurde der Korpus neu, diesmal maschinell, verkodet. Mit Hilfe von Textpack/PC (Züll 1991) wurde eine Häufigkeitsauszählung des Textes vorgenommen. Aus dieser Liste wurden die häufigsten 43 Substantive und Adjektive ausgewählt und als Grundlage für ein Kodeschema genommen. Da der größte Textanteil im Korpus niederländisch ist, wurden die englischen Gegenstücke zu den niederländischen Wörtern mit aufgenommen. Durch Setzen von *wild-cards*<sup>33</sup> wurden auch Pluralbildungen und Kasus berücksichtigt. Dies war

allerdings nicht in allen Fällen sinnvoll möglich: So konnte bei "Mann" nur der Singular aufgenommen werden, da das englische "men" (Männer) sich mit niederländisch "men" (man) überschneidet und zu einer Verzerrung geführt hätte. In der folgenden Liste sind die Kodeetiketten, die ihrer Bildung zugrunde liegende Wortfrequenz, sowie die anschließend im Text markierten Indikatoren aufgeführt.

Wortliste für die formale Analyse

1. "Arbeit" (55) WORK; WERK
2. "Bandung" (50) BANDUNG
3. "Bolkenstein" (50) BOLKENSTEIN\*
4. "Bruder" (65) BROTHER; BROER
5. "Dag" (109) DAY\*; DAG; DAGEN
6. "Dood" (63) DEATH; DOOD
7. "Eltern" (95) PARENTS; OUDERS
8. "Erzählung" (125) STORY; STORIES; VERHAAL; VERHALEN
9. "Eten" (136) EAT; FOOD; ETEN
10. "Fabrik" (86) \*FABRIEK; \*FABRIEKEN; \*FACTORY; \*FACTORIES
11. "Familie" (176) FAMILY; FAMILIE\*
12. "Felix" (84) FELIX
13. "Frau" (119) WOMAN; WOMEN; VROUW; VROUWEN
14. "Gefängnis" (45) GEVANGENIS; JAIL
15. "Geld" (45) GELD; MONEY
16. "Haus" (222) HOUSE; HOUSES; HUIS; HUIZEN
17. "Henriëtte" (107) HENRIËTTE
18. "Holland" (206) HOLLAND\*
19. "Indien" (154) INDIË
20. "Indisch" (302) INDISCH\*
21. "Indonesien" (309) INDONESIA\*; INDONESIAË\*; INDONESIAISCH\*
22. "Jahr" (285) YEAR; YEARS; JAAR; JAREN
23. "Kamp" (79) \*KAMP; \*KAMPEN; \*CAMP; \*CAMPS
24. "Kind" (268) KIND; KINDEREN; CHILD; CHILDREN
25. "Klaas" (54) KLAAS
26. "Kultur" (68) CULTUUR; CULTURE; CULTUREN; CULTURES
27. "Land" (91) LAND; LANDEN; COUNTRY; COUNTRIES
28. "Leben" (125) LEVEN; LIFE
29. "Mann" (141) MAN
30. "Marlene" (82) MARLENE
31. "Menschen" (327) MENSEN; PEOPLE
32. "Michel" (102) MICHEL
33. "Mutter" (347) MOEDER; MOTHER
34. "Nederland" (451) NEDERLAND\*; DUTCH
35. "Oma" (166) OMA
36. "Onkel" (69) OOM; UNCLE
37. "Oorlog" (133) OORLOG; WAR
38. "Schule" (153) SCHOOL; SCHOOLS; SCHOLEN
39. "Surabaya" (77) SURABAYA
40. "Tante" (285) TANTE; AUNT
41. "Vater" (254) VADER; FATHER
42. "Zeit" (206) TIJD; TIME
43. "Zimmer" (55) KAMER; KAMERS; ROOM; ROOMS

33. Platzhaltern, die es ermöglichen, nach einem Wortstamm und seinen möglichen Endungen zu suchen.

Nach diesem Vorgehen konnte eine neue Matrix erstellt werden, die die Informanten als Fälle und die von ihnen benutzten häufigsten Wörter als Variablen abbildet. Natürlich sagt das Auftreten eines Wortes alleine im Text nichts über den Zusammenhang aus, in dem es benutzt wird. Es können, schlimmer noch, gegensätzliche Kontexte sein, in denen ein Wort auftaucht. So machen die üblichen univariaten und bivariaten Auswertungsverfahren zunächst einmal keinen Sinn, da inhaltliche Abbildungen im folgenden nicht überprüft werden.

Tabelle 16 demonstriert die Textbeschreibung durch die häufigsten Wörter.

"N abs (Freq)" bezeichnet die Gesamtzahl der vergebenen Codes aufgrund der häufigsten Wörter, wobei manche Segmente mit mehreren Codes belegt sein können. "N Kat (Freq)" bildet ab, wie viele der zur Verfügung stehenden 43 Wörter in der Darstellung des Informanten auftauchen. "N Tags" gibt die Anzahl der insgesamt produzierten Segmente wieder.

Dabei fällt zunächst auf, daß die Verteilung der Codes nach diesem Verfahren viel homogener erfolgt als im ersten Fall (s. Tabelle 7). Es gibt weniger Ausreißer, und insgesamt sind die unterschiedlichen Textanteile der Informanten viel dichter durch Codes belegt. Pro Segment werden dann auch 6.2 Codes vergeben. Mittelwert, Median und Mode von "N Kat (Freq)" nähern sich an und weisen auf eine symmetrische Verteilung hin.

Das Scatterplot von "N abs (Freq)" und "N Kat (Freq)" (Abb. 13) ist gleichmäßiger als das vergleichbare aus Abbildung 7.

Auch hier wurde in der Abbildung der Informant Fünf aus den gleichen Gründen nicht mit aufgenommen, wohl aber bei der Mittelwert- (24,16) und Regressionsgradenberechnung. Auch diese Art der Darstellung zeigt eine Parallele zur Verkodung durch Kategorien auf: Steigt der Textanteil der Informanten, so ist die "Chance" möglichst alle Codes anzusprechen hoch. Während im ersten Fall nicht einmal Informant Fünf, mit seinem extrem hohen Redeanteil alle Kategorien produzierte, sind es bei der Verkodung durch die häufigsten Wörter immerhin 3 Informanten, die mehr als 40, also über 90% der möglichen Codes erhalten.

Dieser optische Eindruck kann durch Messung nachgewiesen werden. *Pearsons r* zeigt zum Teil dramatisch hohe Ergebnisse (siehe Tabelle 17).

Es gibt einen extremen Zusammenhang zwischen der Gesamtzahl kodierter Segmente auf Basis des Kategoriensystems und der häufigsten Wörter ( $r=.97^{**}$ ). Die inhaltlich gehaltvollen Textsegmente eines

*Tabelle 16: Kodes und Tags auf Grundlage der häufigsten Wörter*

Informanten	N abs (Freq)	N Kat (Freq)	N Tags
INF01 (1)	211	36	183
INF02 (2)	37	18	54
INF03 (3)	14	9	32
INF04 (4)	68	22	137
INF05 (5)	1289	43	2608
INF06 (6)	141	30	151
INF07 (7)	223	35	176
INF08 (8)	138	30	132
INF09 (9)	383	42	187
INF10 (10)	135	35	27
INF11 (11)	10	7	15
INF12 (12)	6	6	3
INF13 (13)	8	6	6
INF15 (14)	227	30	434
INF16 (15)	56	18	22
INF17 (16)	73	23	116
INF18 (17)	174	34	120
INF19 (18)	93	18	62
INF20 (19)	1	1	5
INF21 (20)	446	39	482
INF22 (21)	155	34	37
INF23 (22)	244	39	131
INF24 (23)	126	32	24
INF25 (24)	5	3	5

INF26 (25)	141	29	125
INF27 (26)	142	31	110
INF28 (27)	70	24	155
INF29 (28)	409	41	298
INF30 (29)	52	22	86
INF31 (30)	427	35	357
INF32 (31)	32	17	82
INF33 (32)	19	12	34
INF34 (33)	11	9	19
INF35 (34)	1	1	6
INF36 (35)	95	24	38
INF37 (36)	121	29	32
INF38 (37)	108	31	19
INF39 (38)	118	31	39
INF40 (39)	128	29	120
INF41 (40)	48	21	43
INF42 (41)	107	29	67
INF44 (42)	8	5	23
INF46 (43)	123	29	319
TOTAL 43	6.423	1.039	7.121

Mittelwert                      149,37                      24,16                      165,60

Std. Deviation	212,54	12,09	398,47
Std. Error	32,41	1,84	60,77
Median	108	29	67
Mode	1 <sup>a</sup>	29	5
Varianz	45172,14	146,19	15.8781,77

a. Von mehreren *modes* wird hier nur der Kleinste angezeigt.

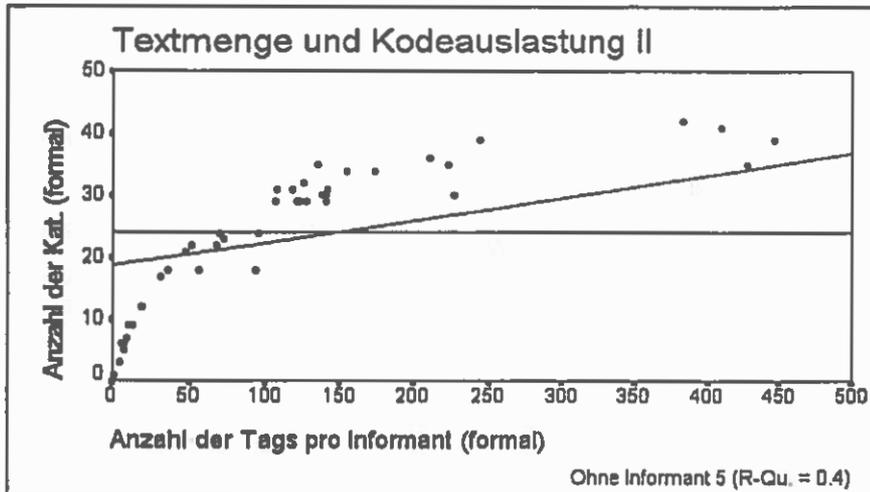


Abbildung 13

**Tabelle 17: Pearsons 'r' für Zusammenhänge zwischen zwei Codesystemen**

	-- Correlation Coefficients --			
	N abs	N Kat	N Tags	N abs (FREQ)
N abs (FREQ)	,9709 ( 43) P= ,000	,8874 ( 43) P= ,000	,9271 ( 43) P= ,000	
N Kat (FREQ)	,5427 ( 43) P= ,000	,8131 ( 43) P= ,000	,3976 ( 43) P= ,008	,6365 ( 43) P= ,000

(Coefficient / (Cases) / 2-tailed Significance)  
 " . " is printed if a coefficient cannot be computed

Informanten lassen sich also rein quantitativ betrachtet, durch beide Verfahren aus dem gesamten Korpus herausfiltern. Auch die Kodeauserschöpfung ("N Kat" und "N Kat (Freq)") korrelierten extrem hoch ( $r = .81^{**}$ ). Der Vergleich der beiden Scatterplots täuscht also nicht: Textmenge und Kodeauslastung hängen bei beiden Verfahren zusammen. Für eine Analyse von Texten ist also die reine Menge, die von einzelnen Informanten produziert wird, von entscheidender Bedeutung. Es ist keineswegs so, daß einzelne Personen in einem mündlichen Text, wie dem vorliegenden, die Dinge "auf den Punkt" bringen könnten, in dem sie also in wenigen Sätzen die gesamten Ideen produzierten. Unabhängig von der Unterschiedlichkeit der Informanten qua sozialem Hintergrund und persönlichen Faktoren muß also eine bestimmte Textmenge vorliegen, damit die für die Fragestellung relevanten Themen angehend ausschöpfend



Folgende Informantengruppen wurden gebildet:

- A: Richie (20; 19<sup>34</sup>), Andrew (25; 24)
  - B: George (03; 03), Rein (44; 42)
  - C: Woody (12; 12), Laila (13; 13), Jane (11; 11), Didier (32; 31)
  - D: Lilian (34; 33), Erna (41; 39), Ellie (16; 15), Asta (33; 32), Meike (19; 18), Emiel (30; 29), Cleo (46; 43)
  - E: Harko (17; 16), Eveline (02; 02), Henry (04; 04)
  - F: Astrid (36; 35), Antje (26; 25), Ruud (27; 26), Edmond (37; 36), Tania (38; 37), Gabriella (24; 23), Margriete (18; 17), Olga (39; 38)
  - G: Toni (28; 27), Anna (40; 39), Henriëtte (42; 41)
  - H: Isabella (15; 14), Michel (10; 10), Pia (22; 21), Oscar (23; 22), Simone (31; 30), Sjakkie (21; 20), Saskia (29; 28), Jaqueline (09; 09), Marlene (05; 05)
  - I: Ivo (06; 06), Klaas (01; 01), Kurt (07; 07), Gerard (08; 08)
- Isoliert: Gaston (35; 34)

Wie erwähnt, ergibt sich die Nähe der Informanten nunmehr nicht mehr aus ihren Ideen, sondern aus der formalen Struktur ihrer Sprache. Ein Vergleich mit Abbildung 11 zeigt aber, daß 26 von 43 Informanten auch hier in Beziehung stehen<sup>35</sup>. Mit anderen Worten, 60,5% der Informanten werden nur aufgrund der Struktur ihrer Sprache genauso verbunden, wie über ihre explizit formulierten Ideen. Und dies bei einem im Hinblick auf dieses Verfahren nicht optimal erhobenen Textkorpus, der sich durch starke Heterogenität in den Textanteilen sowie durch geringe Moderation auszeichnet. Ob Informanten in andere Gruppen "rutschen", scheint stark von der produzierten Textmenge abzuhängen. Diese Instabilität kann also leicht abgestellt werden.

Interessant ist nun ein Blick auf die Cluster der häufigsten Wörter (vergleiche Abbildung 15).

Folgende Gruppen wurden gebildet:

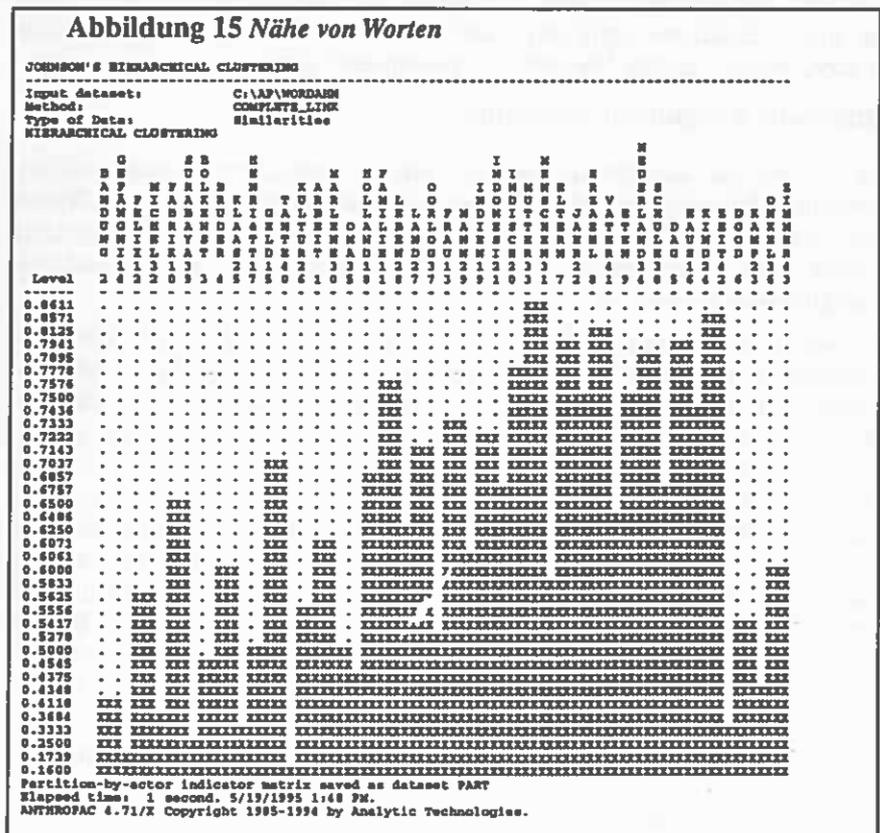
- A: Bandung, Gefängnis
- B: Felix, Michel, Fabrik, Surabaya
- C: Bolkenstein, Bruder, Klaas, Henriëtte, Geld, Tante

34. Wieder entspricht die erste Zahl in Klammern dem im Korpus verwendeten Informantenetikett, die zweite bezieht sich auf die in den folgenden Analysen verwandte Fallzahl.

35. Woody und Laila; Lilian, Ellie und Asta; Astrid, Antje, Ruud, Edmond, Gabriella, Margriete und Olga; Anna und Henriëtte; Michel, Simone, Saskia, Jaqueline und Marlene; Pia, Oscar und Sjakkie; Ivo, Kurt und Gerard; Gaston.

- D: Kultur, Arbeit, Marlene, Oma
- E: Holland, Familie, Leben, Land, Krieg
- F: Frau, Mann, Indien, Indonesien, Indisch, Mutter, Menschen
- G: Eltern, Jahr, Erzählung, Vater
- H: Essen, Niederlande, Schule
- I: Tag, Haus, Kind, Zeit
- J: Tod, Lager, Onkel, Zimmer

Diese Konfiguration reizt geradezu zur Interpretation, wiewohl auch hier



beachtet werden muß, daß es sich um statistische Assoziationen handelt, über deren inhaltliche Verbindung keine Aussage gemacht werden kann. Es ist aber schon frappierend, daß vor dem Hintergrund der bisher gemachten Ergebnisse, diese Assoziationen eben auch inhaltlich einen Sinn machen. Orte sind mit bestimmten Wörtern gekoppelt, die sich analog in den

Themen wiederfinden: *Bandung* als Mittelpunkt traumatischer Lagererfahrung; *Surabaya* mit *Felix* und *Michel* und dem für die Familie wichtigen Unternehmen. *Kultur* ist eng mit *Marlene* und *Oma* gekoppelt, aus dem Korpus weiß man, daß in fast allen Fällen mit "Oma" *Anna* gemeint ist. Also wird auch hier aus der räumlichen Assoziation die inhaltliche erkennbar.

*Krieg* findet sich überraschenderweise in einem Cluster mit *Holland*, während *Indonesien*, *indisch* und *Indien* mit Bezeichnungen für *Menschen* und Personen verbunden sind. *Erzählung*, auch dies nicht erstaunlich, steht im engen Zusammenhang mit *Vater*, während kulturelle Basisfunktionen (*Essen*, *Schule*) an die *Niederlande* gekoppelt<sup>36</sup> sind.

#### Quadratic Assignment Procedure

Es scheint also im Hinblick auf die Informanten einen Zusammenhang zwischen der strukturellen Abbildungsmöglichkeit der häufigsten Wörter und den inhaltlichen Kategorien zu geben. Man kann dies überprüfen, indem man Ähnlichkeitsmatrizen miteinander korreliert (Quadratic Assignment Procedure).

Zur Berechnung des Ähnlichkeitsmaßes wurden zwei verschiedene Prozeduren angewandt. Die beiden Ausgangsmatrizen (Informanten und Kategorien sowie Informanten und häufigste Wörter) wurden dichotomisiert. Es wurde also immer nur das Auftreten eines Ereignisses gewertet. Danach wurde im ersten Fall die Matrix gespiegelt und eine Ähnlichkeit über Summenbildung geprüft. Dadurch entstehen zwei Informanten-Informanten-Matrizen, die für jedes Paar von Informanten die absolute Zahl der übereinstimmenden Kategorien einerseits und im anderen Fall die absolute Zahl der übereinstimmenden Worte enthält. Diese Art der Ähnlichkeit kann man als "weighted similarity" (Borgatti 1992b: 62) begreifen. Bei der Quadratic Assignment Procedure wird für jede korrespondierende Zelle der beiden Matrizen *Pearsons r* berechnet. Danach wird eine der beiden Matrizen nach einem Zufallsverfahren mit Zahlen gefüllt und danach die Korrelation erneut berechnet. Dies geschah im folgenden Fall tausendmal. Mit dem zweiten Schritt wird somit überprüft, ob die Korrelation der unveränderten Ausgangsmatrizen zufällig ist oder nicht (Borgatti 1992b: 85). Im vorliegenden Fall ergab sich ein überaus starker Zusammenhang:

---

36. Hier kann allerdings auch eine systematische Verzerrung vorliegen, weil mit der Kette NEDERLAND\* z.B. nicht zwischen *Niederland* und *Indische Niederlanders* differenziert wird.

Abbildung 16 Qap Ergebnisse

QAP MATRIX CORRELATION		
# of Permutations:	1000	
	CORRELATION	MATCHES
Observed value:	0.784	0.064
Average:	-0.001	0.041
Standard deviation:	0.126	0.010
Proportion as large:	0.000	0.020
Proportion as small:	1.000	0.983
Hubert's gamma: 158908.000		
Elapsed time: 20 seconds. 5/20/1995 2:22 PM.		
ANTHROPAC 4.71/X Copyright 1985-1994 by Analytic Technologies.		

Bei tausend zufälligen Permutationen stellt sich keine ähnliche Verteilung wie durch den empirischen Befund ein. Die beiden Matrizen bilden also in hohem Maße ( $r=.78$ ) (Abb. 16) den gleichen Sachverhalt ab.

Konservativeres Vorgehen verlangt jedoch eine Bestimmung der Ähnlichkeit durch *positiv matches*, also den Tanimoto-Koeffizienten. Dieser bezieht, wie oben aufgezeigt, auch die individuellen Ausprägungen mit in die Berechnung der Ähnlichkeit ein. Aber auch in diesem Fall ist der Zusammenhang mit  $r=.7$  (Abb. 17) überzeugend stark:

Abbildung 17 Qap Ergebnisse

QAP MATRIX CORRELATION		
# of Permutations:	1000	
	CORRELATION	MATCHES
Observed value:	0.703	0.030
Average:	0.001	0.009
Standard deviation:	0.096	0.006
Proportion as large:	0.000	0.011
Proportion as small:	1.000	0.994
Hubert's gamma: 116.905		
Elapsed time: 20 seconds. 5/20/1995 2:21 PM.		
ANTHROPAC 4.71/X Copyright 1985-1994 by Analytic Technologies.		

Die Verteilung erfolgt in keinem von 1000 Fällen einem zufälligen Muster.

Dies bedeutet also: Wenn in der einen Matrix ein Paar von Informanten Ähnlichkeit dadurch aufweist, daß eine bestimmte Kategorie in seiner Darstellung in einer bestimmten Häufigkeit vorkommt, so wird dies auch in der anderen Matrix im Hinblick auf die benutzten Wörter der Fall sein - und dieses passiert, absolut oder relativ, auch mit allen anderen Informantenpaaren.

Zwar macht dies nicht die anderen statistischen Verfahren, und vor allem nicht eine qualitative Analyse, überflüssig, ohne die über die Inhalte und spezifischen Kontexte nichts ausgesagt werden kann. Wenn diese Ergebnisse mit anderen Datensätzen repliziert werden können, sind eine Reihe von Konsequenzen denkbar. Zum einen ist der Beweis führbar, daß genügend große Datensätze, was die Abbildung der Informanten angeht,

hinreichend durch die Verteilung der häufigsten Wörter beschrieben werden können. Dadurch wäre es in Zukunft möglich, beliebig große Textmengen im Hinblick auf spezielle Fragestellungen zu analysieren. Zum anderen wäre ein Verfahren gefunden, das im hohen Maße reliabel ist und die systematische Auswertung fremdsprachiger Texte ermöglicht, bei deren Verkodung bisher auf vom Forscher unabhängige Koder verzichtet werden mußte. Eine solche Fragestellung könnte die Messung von Beziehungen der Informanten durch die Struktur ihrer Worte betreffen - und beispielsweise netzwerkanalytische Verfahren einsetzen.

### **Zusammenfassung der quantitativen Analyse**

In der quantitativen Analyse wurde erstens versucht, die Ergebnisse der qualitativen zu bestätigen. Zweitens ging es um den Nachweis, daß bestimmte Aussagen über Text und Informanten auf anderem Wege als über inhaltliche Verkodung möglich sind.

Das am Text entwickelte Kategoriensystem erwies sich als griffig und im Rahmen der Fragestellung gültig. Die Gruppendiskussion als Erhebungsinstrument war insgesamt produktiv. Vorschläge für eine Verbesserung der Erhebungssituation wären eine stärkere Moderation, um mengenmäßig ausgeglichene Texte zu erzeugen.

Aufgrund der niedrigen Fallzahl fand eine Datenreduktion und dadurch ein Informationsverlust statt. Bei 15 von 26 gebildeten Themengruppen konnten keine Zusammenhänge mit externen Variablen festgestellt werden. Die Themen dieser Gruppen wurden also unabhängig von diesen Faktoren übertragen. In den restlichen 11 Gruppen ergaben sich 17 Zusammenhänge. Dabei bestätigten sich im wesentlichen die Ergebnisse der qualitativen Analyse: Ausgerechnet die Elemente, die bei den Älteren und denjenigen, die eine persönliche Erfahrung im Ursprungsland haben, so wichtig für die Aufrechterhaltung eines Gruppenbewußtseins sind, werden nicht übertragen. Zwar fließt wesentliches Wissen, das zum Verständnis des Phänomens Indische Niederländer nötig ist, ungehindert zu den Jüngsten durch, es bleibt aber ohne Bezug zur eigenen Persönlichkeit und führt nicht zur ethnischen Anbindung. Nur ein geringer Teil der Informanten (Gruppe F aus Abbildung 11) bekennt sich uneingeschränkt zur Vergangenheit und zur Gruppe.

Die Clusteranalyse der Themen macht unter anderem die bei der Gruppendiskussion auftretenden Intervieweffekte transparent. Diese führten allerdings nicht zu artefaktischen Trübungen.

Im vorliegenden Fall konnten hohe Zusammenhänge zwischen der Abbildungsfähigkeit von Strukturen des Textes und der Ähnlichkeit der Informanten durch reine Worthäufigkeiten aufgezeigt werden. Dieser Effekt mag auf die beiden optimal ausgearbeiteten Codesysteme zurückgeführt werden. Die Ergebnisse ermutigen - falls sie an anderen Datensätzen replizierbar sind - zum verstärkten Einsatz formaler quantitativer Auswertungsinstrumente bei großen Textmengen und hohen Informantenzahlen.



Im vorliegenden Fall konnten hohe Zusammenhänge zwischen der Abbildungsfähigkeit von Strukturen des Textes und der Ähnlichkeit der Informanten durch reine Worthäufigkeiten aufgezeigt werden. Dieser Effekt mag auf die beiden optimal ausgearbeiteten Codesysteme zurückgeführt werden. Die Ergebnisse ermutigen - falls sie an anderen Datensätzen replizierbar sind - zum verstärkten Einsatz formaler quantitativer Auswertungsinstrumente bei großen Textmengen und hohen Informantenzahlen.

