

# Anhang

## Was ist TEXTPACK?

*Cornelia Züll und Peter Ph. Mohler*

TEXTPACK ist zwar kein Dinosaurier wie sein Bruder der General Inquirer (Züll/Weber/Mohler 1989), aber zum Minidrachen der computerunterstützten Inhalts- und Textanalyse ist das Programm inzwischen ausgewachsen. Auf seine Vergangenheit soll hier, obwohl sie das Licht des Tages nicht scheuen muß, nicht weiter eingegangen werden, sondern es sollen einige Erläuterungen zu dem Programm, über das in den Beiträgen dieses Bandes das eine oder andere gesagt wurde, systematisch zusammengetragen werden.

Mit TEXTPACK stehen in einem einheitlichen Programmsystem sämtliche Standardfunktionen für die Deskription und Exploration von Texten zur Verfügung. Die Besonderheit von TEXTPACK liegt in der Verknüpfung dieser deskriptiven und explorativen Verfahren mit quantifizierenden Verfahren der computerunterstützten Inhaltsanalyse (cui).

Wenn im Zusammenhang mit TEXTPACK über Textdeskription und Exploration gesprochen wird, dann sind darunter insbesondere die komfortablen Häufigkeitslisten der Wörter eines Textes zu verstehen, die nach verschiedensten Kriterien, auf- oder absteigend nach Alphabet oder Häufigkeiten, sortiert werden können. Aber auch die KeyWord-In-Kontext-Prozedur (KWIC), die u.a. erlaubt, Schlüsselwörter in unterschiedlich langem Kontext auszugeben, ist ein wichtiges Standardinstrument der Textexploration. Ganz wichtig in diesem Zusammenhang ist auch die Möglichkeit, den Wortbestand zweier Texte, etwa das Vokabular zweier Redner, direkt miteinander zu vergleichen. Selbstverständlich kann man mit TEXTPACK auch den Text oder Teile davon ausdrucken und einen Index der Textwörter erstellen.

Die Besonderheit von TEXTPACK gegenüber anderen Programmen für die Textanalyse ist die Verknüpfung der Textdeskription und -exploration mit einer algorithmisierten Vercodung des Textes. So können Wörter zu Listen (Kategorien eines inhaltsanalytischen Diktionärs) zusammengefaßt werden, wie z.B. alle Verben der Bewegung, alle Bundeskanzler, alle Wörter, die mit Buchstabenfolge Demokrat... beginnen usw. Die Vercodungsprozedur von TEXTPACK prüft dann, ob eine solche Kategorie in einem Textstück erscheint. Wenn ja, wird entweder gezählt, wie oft diese Kategorie in diesem Textstück erscheint, oder es werden die Abfolgen der Kategorien im Text festgehalten. Das Ergebnis sind dann entweder Kategorienhäufigkeiten oder quasi Zeitreihen, die das Auftreten von Kategorien bezeichnen. Häufigkeiten und Zeitreihen können mit Hilfe von Statistikprogrammen wie SPSS, SAS, NSDstat+, Systat etc. ausgewertet werden.

Damit bei der Vercodung nicht einfach Wörter in Ziffern umgewandelt werden und man nachher nicht so genau weiß, wie das denn so zustande gekommen ist, bietet TEXTPACK eine Reihe von Möglichkeiten der Validierung des Kategorienschematas. So kann man mit TEXTPACK z.B. die Zifferncodes in den Text zurückschreiben, um so die Vercodung überprüfen zu können. Andere Prüftechniken, auf die hier nicht eingegangen werden soll, erleichtern weitere Kontrollen der automatischen Vercodung.

Das ganze läuft auf MS-DOS PCs unter einer eigenen kleinen Benutzeroberfläche (Menü) und auf Großrechnern, die für sehr große Texte auch heute noch hilfreich sein können. Der Textumfang ist in TEXTPACK faktisch unbeschränkt (es können beliebig große Texte (nur begrenzt durch den verfügbaren PC-Plattenspeicherplatz) in einer TEXTPACK-Systemdatei gespeichert und bis zu 999 Kategorien in einem Durchlauf vercodet werden). TEXTPACK akzeptiert ASCII-Zeichen aller Art und kann auch seine Systemdateien in ASCII-formatierte Dateien zurückwandeln.

Die besondere Kombination von Exploration, Deskription und Quantifizierung macht TEXTPACK PC für mehrere große Anwendungsbereiche geeignet: für eine quantitativ ausgerichtete Inhaltsanalyse, für eine eher linguistisch ausgerichtete Textanalyse, für das Datenmanagement in qualitativ ausgerichteten Studien und vieles andere mehr.

Innerhalb der cui unterstützt TEXTPACK sowohl den Ansatz des General Inquirer, wie auch andere Ansätze, etwa von Iker.

Die oben genannten verschiedenen Funktionen von TEXTPACK sind in folgenden Prozeduren zusammengefaßt:

- *SENTENCE* überführt Rohtexte (ASCII-Format) in eine TEXTPACK-Systemdatei. Die Prozedur erlaubt drei unterschiedliche Eingabeformate.

- *SENMERGE* vereinfacht die Ergänzung und Korrektur bestehender Systemdateien.
- *LISTSPLIT* erzeugt eine Datei von Einzelwörtern (Vertikaltext) und druckt gegebenenfalls den Text aus (zwei Druckformate).
- *FREQ* zählt Worthäufigkeiten über den gesamten Text bzw. dessen Untereinheiten aus und berechnet den Type-Token-Ratio (TTR). Die Datei der Worthäufigkeiten kann durch verschiedene Optionen in ihrem Umfang eingeschränkt werden.
- *KWIC* stellt einzelne Wörter, Wortanfänge oder Mehrwortverbindungen in ihrem Kontext in der Form "keyword in context" oder "keyword out of context" dar.
- *XREF* erzeugt einen Index aller Wörter eines Textes oder eines Subsets der Wörter des Textes (Konkordanzen).
- *WORDCOMP* vergleicht den Wortschatz zweier Texte (verschiedene Ausgabeformate).
- *TAGCODER* vercodet automatisch einen Text mit Hilfe eines vorgegebenen Kategorienschemas (inhaltsanalytisches Diktionär). Die Prozedur erzeugt sowohl Dateien mit den Kategorienhäufigkeiten als auch mit den Kategorienabfolgen. Die Validität der Vercodung kann mit mehreren Optionen überprüft werden.
- *SUBSEL* wählt Texteinheiten aus einer Textdatei aus. Basis für die Auswahl sind numerischen Informationen in einer zweiten Datei, die als portable SPSS-Datei (EXPORT) oder eine Rohdaten-ASCII-Datei vorliegen kann.
- *REFORM* setzt eine Systemdatei in Rohtextformat für die Weitergabe bzw. Eingabe in andere Programme um.

Weitere Informationen und das System selbst sind beim Zentrum für Umfragen, Methoden und Analysen in Mannheim erhältlich. Ein ausführliches deutsches Handbuch wurde von C. Züll, P. Ph. Mohler und A. Geis (1991) unter dem Titel "Computerunterstützte Inhaltsanalyse" veröffentlicht.

## Literatur

- Züll, C., Weber, R.P., Mohler, P.Ph. (1989). Computer-Aided Text Classification for the Social Sciences: The General Inquirer III. Mannheim: ZUMA.
- Züll, C., Mohler, P. Ph., Geis, A. (1991). Computerunterstützte Inhaltsanalyse mit TEXTPACK PC. Stuttgart, New York: Fischer.

# Aus dem Programm Sozialwissenschaften

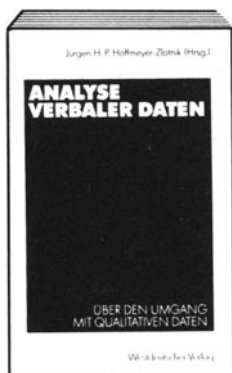


Gerhard Arminger und  
Franz Müller

## **Lineare Modelle zur Analyse von Paneldaten**

1990. 228 S. Kart.  
ISBN 3-531-12176-6

Dieses Buch ist eine Einführung in lineare Modelle zur Analyse von Paneldaten. Paneldaten entstehen, wenn Stichproben über mehrere Zeitpunkte hinweg untersucht werden. Sie werfen spezielle Probleme durch die Berücksichtigung individueller und zeitlicher Variation auf. Alle Modelle werden anhand eines Beispieldatensatzes mit beigefügten LISREL- und LISCOMP-Programmen dargestellt und analysiert.



Jürgen H. P. Hoffmeyer-  
Zlotnik (Hrsg.)

## **Analyse verbaler Daten**

Über den Umgang mit  
qualitativen Daten.

1992. VIII, 424 S. Kart.  
ISBN 3-531-12360-2

Dieses Buch behandelt die unterschiedlichen Möglichkeiten und Ansätze der Analyse qualitativer Daten. Hierbei wird besonderer Wert darauf gelegt, daß die unterschiedlichen Analyseansätze nicht nur methodologisch diskutiert, sondern auch hinsichtlich methodischer Anwendungen an jeweils konkreten Projektfragestellungen und -daten demonstriert werden. Damit ist dieser Band als eine praxisorientierte Einführung in die Analyse verbaler Daten zu sehen.



Klaus G. Troitzsch

## **Modellbildung und Simulation in den Sozialwissenschaften**

1990. 206 S. Kart.  
ISBN 3-531-12150-2

Das Buch – hervorgegangen aus einer über mehrere Jahre hinweg immer wieder gehaltenen Vorlesung – gibt anhand von Beispielen eine Einführung in die wichtigsten Verfahren der mathematischen und computergestützten Modell- und Theoriebildung in den Sozialwissenschaften. Behandelt werden deterministische und stochastische Modelle von Prozessen der individuellen Meinungsbildung und der Veränderung von Einstellungen in größeren Kollektiven.



WESTDEUTSCHER  
VERLAG

OPLADEN · WIESBADEN