

Die technikunterstützte Analyse von qualitativen Daten mit Word¹

Bruno Nideröst

Keywords: Code and Retrieve, Sortierung, Häufigkeitsauszählung, technikunterstützte Analyse, qualitative Daten, Word, Textverarbeitung

Zusammenfassung: Dass sich Textverarbeitungsprogramme auch für die technikunterstützte qualitative Datenanalyse (QDA) einsetzen lassen, ist schon von verschiedenen Anwendern beschrieben worden (FIELDING & LEE 1998). Die entsprechenden Möglichkeiten von Word werden jedoch häufig unterschätzt (KUCKARTZ 1999). Dank verschiedener Standardbefehle sind zahlreiche elementare Funktionen der technikunterstützten Analyse wie Sortierungen, Code-and-Retrieve und Häufigkeitszählungen möglich. Das Programm ist besonders für jene Einsteiger in die technikunterstützte Analyse interessant, welche erste Erfahrungen sammeln wollen, bevor sie die Kosten und Mühen der Anschaffung spezifischer QDA-Software auf sich nehmen. Word kommt auch jenen qualitativen Forschern entgegen, welche die Vorteile der elektronischen Textverarbeitung gerne nutzen, der technikunterstützten Datenanalyse gegenüber jedoch skeptisch eingestellt sind (KELLE 1995), denn mit Word lässt sich ein fließender Übergang von manuellen zu technikunterstützten Auswertungsmethoden bewerkstelligen. Für routinierte QDA-Anwender kann ein gelegentlicher Einsatz von Word Sinn machen, falls die Formatierung und der Import der Daten in die gewohnte QDA-Software Probleme bereitet oder bei der Auswertung jene einfachen Verfahren angewandt werden sollen, die Word sehr effizient übernehmen kann. Der folgende Artikel befasst sich mit den wichtigsten Voraussetzungen und Möglichkeiten von Word für die computerunterstützte Analyse qualitativer Daten.

Inhaltsverzeichnis

- [1. Die Datenerfassung in einer Basisdatei](#)
- [2. Die Datensortierung in der Basistabelle](#)
- [3. Das breite Anwendungsspektrum des Such- und Ersetz-Befehls](#)
- [4. Die Kombination von verschiedenen Auswertungsoptionen](#)
- [5. Fazit](#)

[Danksagung](#)

[Literatur](#)

[Zum Autor](#)

[Zitation](#)

1. Die Datenerfassung in einer Basisdatei

Wenn qualitative Daten mit Hilfe des Computers ausgewertet werden sollen, so werden sie in der Regel zuerst in einem Textverarbeitungsprogramm in ein passendes Format gebracht und anschließend in ein spezifisches QDA-Programm importiert. Dieses Formatieren und Importieren entfällt, wenn man die Analyse direkt mit jenem Textverarbeitungsprogramm vornimmt, in welchem die Daten gespeichert wurden. Anders als bei der Verwendung von spezifischer Software für die Qualitative Datenanalyse kann mit Word jederzeit an den Daten gearbeitet werden, so dass Schriftgrößen, Schriftarten und Sonderzeichen in jeder

1 Die technischen Angaben beruhen auf der Programmversion Word 2000 von Microsoft.

Etappe des Analyseprozesses an spezifische Bedürfnisse angepasst werden können. [1]

Die technikerunterstützte Analyse mit Word setzt voraus, dass alle gemeinsam zu analysierenden Daten in derselben Datei gespeichert werden. Das Handling dieser mitunter umfangreichen Basisdatei muss nicht schwerfälliger sein, als die übliche Form der separaten Speicherung von Interviewprotokollen und Dokumenten; wer zum Beispiel ein bestimmtes Interviewprotokoll suchen will, findet es dank dem Word-Suchbefehl unter *Bearbeiten / Suchen* rasch. [2]

Eine der wichtigsten Analysemöglichkeiten besteht in der automatischen Tabellierung und Sortierung aller Textsegmente, worauf im nächsten Kapitel noch genauer eingegangen wird. Word orientiert sich bei dieser automatischen Einreihung der Textsegmente in Tabellenfelder an den Absatzmarken innerhalb der Basisdatei². Dieser Umstand wird am einfachsten bereits bei der Transkription der Primärdaten berücksichtigt, indem für thematisch unterschiedliche Textsequenzen bewusst separate Abschnitte gebildet werden. Zwingend erforderlich sind jedoch weder das laufende Speichern aller Daten in einer einzigen Basisdatei noch das frühzeitige Einfügen von geeigneten Absatzmarken für die Festlegung von Texteinheiten. Diese beiden Schritte lassen sich auch erst nach abgeschlossener Datensammlung und -erfassung nachholen, falls man beispielsweise nicht von den gewohnten Darstellungs- und Speicherroutinen abweichen mag oder erst im Verlauf der Datenerfassung beschließt, die Auswertungen technikerunterstützt vorzunehmen. [3]

2. Die Datensortierung in der Basistabelle

Das Tabellieren und Sortieren der Daten ist eine der wichtigsten Möglichkeiten, welche Word für die technikerunterstützte Analyse zur Verfügung stellt. Am einfachsten lässt sie sich am Beispiel der Auswertung von Fragebogeninterviews erläutern. Diese zeichnen sich dadurch aus, dass die Fragen – anders als bei offeneren Leitfaden-Interviews – in fixer Reihenfolge gestellt werden und der Interviewer nicht mit spontan formulierten Zusatzfragen nachhakt. Für eine systematische Auswertung der Protokolle ist es dann meist sinnvoll, sämtliche Antworten zu einer bestimmten Frage aufzulisten, um die Aussagen der verschiedenen Interviewpartner vergleichen zu können. Mit Word lässt sich eine solche Antwortliste leicht herstellen. Nachdem alle Interviewprotokolle in einer Basisdatei gesammelt und Absatzmarken vor jeder neuen Frage eingegeben wurden, lässt sich die markierte Basisdatei mit *Tabelle / Umwandeln / Text in Tabelle* in eine einspaltige Tabelle transformieren. Word reiht automatisch jede Texteinheit nach einer Absatzmarke in ein separates Tabellenfeld ein; wenn die Absatzmarken also in der beschriebenen Weise eingegeben wurden, ist in jedem Tabellenfeld je eine Frage und die dazugehörige Antwort enthalten. Die einfachste Option besteht nun darin, die Basistabelle zu markieren und mit *Tabelle / Sortieren* zu ordnen. Da sich Word bei der Sortierung an den führenden Zeichen in jedem Tabellenfeld orientiert, ist die gewünschte Ordnung aller

² Word kann die Zelleneinteilung auch anhand von benutzerdefinierten Zeichen vornehmen. Absatzmarken heben die verschiedenen Textstellen jedoch optisch besser voneinander ab.

Antworten zu einer bestimmten Frage damit bereits erreicht, denn die Frageformulierungen funktionieren als Sortierkriterium. Ganz ähnlich ließe sich übrigens auch das Protokoll einer Gruppendiskussion gemäß den Gesprächsbeiträgen pro Sprecher sortieren, sofern die Namen der Sprechenden ihren jeweiligen Aussagen vorangestellt werden und vor den Namen eine Absatzmarke eingefügt wird. Das Sortierergebnis wäre in diesem Falle eine Liste aller Wortbeiträge pro Gruppengesprächsteilnehmer. [4]

Erfreulich ist in beiden Fällen der minimale Aufwand, mit dem das angestrebte Ziel erreicht wird. Je nach gewünschter Analysestrategie fehlen beim Sortierergebnis jedoch Hinweise auf den Diskussionszusammenhang und wichtige Fallvariablen (wie beispielsweise Geschlecht, Alter und professionelle Rolle der verschiedenen Gesprächsteilnehmer bzw. Interviewpartner u.ä.). Zudem ist bei den meisten Forschungsprojekten eine feinkörnigere Auswertung erforderlich, welche sich an thematischen Textsequenzen orientiert und nicht an den strukturell vorgegebenen Texteinheiten von Interviewfragen oder Diskussionsbeiträgen in einem Gruppengespräch. Die Tabellierung der Basisdatei mit anschließender Sortierung lässt sich auf die angesprochenen Analyseprobleme anwenden, indem das bereits vorgestellte Vorgehen ausgebaut und verfeinert wird. In den Protokollen von Leitfadeninterviews wird zu diesem Zweck nach jeder eigenständigen Themensequenz eine Absatzmarke eingefügt, bevor sie zusammen mit der Basisdatei in eine Tabelle umgewandelt wird. Nun wird die resultierende Tabelle nicht sofort sortiert, sondern erst um weitere Spalten ergänzt, welche der detaillierten Ordnung der einzelnen Textsegmente dienen. Die Basistabelle setzt sich nach diesem Arbeitsschritt aus der Primärdatenspalte und den verschiedenen Sortierspalten zusammen. In die einzelnen Felder der Sortierspalten lassen sich nun zu jedem Textsegment passende Sortierkürzel eingeben. Je nach Fragestellung und Auswertungsstrategie kommen Sortierspalten in Frage für

- die Fallvariablen (das Alter und Geschlecht der Befragten, ihre berufliche Position oder die von ihnen repräsentierte Institution u.ä.)
- die Interviewnummer
- einen Laufindex für die einzelnen Textsegmente (wegen der mit der späteren Sortierung verbundenen Dekontextualisierung der einzelnen Textsegmente ist ein solcher Laufindex sinnvoll, um mit einem entsprechenden Sortierbefehl jederzeit die ursprüngliche Ordnung der Basistabelle wieder herstellen zu können)
- die Codeliste
- chronologische Angaben (im Format JJMMTT; dabei entspricht J = Jahr, M = Monat und T = Tag. Entsprechende Sortierungen können beispielsweise für den Aufbau einer Fallchronologie oder für die Periodisierung einer Diskursanalyse sinnvoll sein.) [5]

Die einzelnen Felder dieser Sortierspalten sind gemäß den dazugehörigen Textsegmenten mit Sortierkürzeln zu beschriften. Gleichbleibende Sortierkürzel (wie Interviewnummern und Fallvariablen) lassen sich dabei mit Kopierbefehlen

rasch erzeugen, wenn man die gesamte Basistabelle vorübergehend in Excel bearbeitet. Insbesondere können Laufindizes in Excel automatisch erstellt werden (*Bearbeiten / Ausfüllen / Reihe / Spalten / Linear*). Andere Sortierspaltenfelder – wie etwa jene für Codes und Datumsangaben – sind einzeln zu beschriften, was den mit Abstand aufwändigsten Arbeitsschritt darstellt; er entspricht dem manuellen Kodieren, wie man es von QDA-Programmen gewohnt ist. [6]

Nach der Beschriftung der Sortierspalten lässt sich die gesamte Basisdatei gemäß den verschiedenen Sortierkürzeln ordnen, wobei Word grundsätzlich nach drei Schlüsseln gleichzeitig sortieren kann. Beispielsweise lassen sich die Daten nach 1. Codes, 2. Beruf und 3. Geschlecht ordnen. Die Anzahl Sortierschlüssel lässt sich indirekt noch erhöhen, in dem beispielsweise die Angaben für den Beruf (mit *L* als Sortierkürzel für Lehrperson) und für das Geschlecht (*w* oder *m*) in einer einzigen Sortierspalte zusammengefasst werden (Sortierkürzel *Lw* oder *Lm* usw.); ein Sortierbefehl auf diese Spalte ordnet die Daten dann prioritär nach Beruf und innerhalb derselben Berufsgruppe nach Geschlecht. Eine zweite Erweiterungsmöglichkeit zu den drei Sortierschlüssel von Word besteht darin, eine erste Sortierung der Basistabelle – beispielsweise gemäß der Spalte Alter – durchzuführen und die gesamte Basistabelle anschließend in separate Altersklassen aufzuteilen, welche mit den drei Sortierkriterien von Word noch detaillierter geordnet werden. Mit diesen Mitteln lassen sich je nach Bedarf sehr differenzierte Datenordnungen realisieren. [7]

Die Sortierspalte für die Codes dürfte für viele Auswertungsbedürfnisse besonders wichtig sein, denn das Ordnen der Daten nach Codes ist häufig eine aussichtsreichere Strategie als das eingangs vorgestellte einfache Sortieren nach Fragen. Dank geschickt festgelegten Codes lässt sich mit der Basistabelle ein sogenanntes kontrastierendes Retrieval (KUCKARTZ 1999) durchführen. Dabei sollen ganz bestimmte Antworten eines Befragten miteinander verglichen werden, um allfällige Diskrepanzen zwischen dessen Aussagen sichtbar zu machen. Was hält ein Befragter beispielsweise für machbar, und wie sehen im Vergleich dazu seine Wünsche aus? Wenn die entsprechenden Textstellen aus den Interviews einander gegenüber gestellt werden, lässt dies interessante Rückschlüsse zu. Die für ein solch kontrastierendes Retrieval erforderliche Analyselogik lässt sich mit der Basistabelle umsetzen, indem die Codes für *Wunsch* und *Machbarkeitsvorstellung* nach der Sortierung unmittelbar nacheinander zu liegen kommen. Das lässt sich erreichen, indem diesen beiden Sortierkürzeln beispielsweise je eine geeignete Zahl vorangestellt wird, also etwa *1Wunsch* und *2Machbarkeitsvorstellung*. Nun lassen sich mit einem Sortierbefehl mit den Schlüsseln 1. *Interviewnummer* und 2. *Codes* die Machbarkeitsvorstellungen jedes Befragten mit seinen Wunschvorstellungen kontrastieren. [8]

Die Tabellierung und Sortierung der Basisdatei eröffnet zahlreiche Analysemöglichkeiten, wie ein weiteres Beispiel illustriert. Bei vielen Forschungsprojekten besteht ein Interesse an einer Tabelle zu Alter, Geschlecht und weiteren Fallvariablen aller Befragten. Eine solche Tabelle lässt sich bequem herstellen, indem bei der Transkription der Interviews die jeweiligen Fallvariablen in eine separate Zeile des Protokolls getippt werden. Bei der Umwandlung der

Basisdatei in eine Basistabelle werden diese Zeilen – wie alle anderen durch Absatzmarken getrennten Textsegmente – in separate Tabellenfelder eingeteilt. In der Basistabelle schließlich setzt man zu jeder Zeile mit Fallvariablen einen entsprechenden Code in eine geeignete Sortierspalte, wonach ein Sortierbefehl auf diese Spalte alle Falldaten zusammenstellt. [9]

Die genannten Beispiele sollen die verschiedenen Möglichkeiten der Basistabelle lediglich etwas verdeutlichen, ohne sie jedoch abschließend zu behandeln, denn mit der tabellarischen Sortierfunktion lassen sich sehr vielfältige Analysebedürfnisse abdecken. Es hängt von der jeweiligen Fragestellung und dem konkreten Forschungsprojekt ab, nach welchen Konventionen die Basisdatei in einzelne Tabellenfelder zerlegt wird, wie viele und welche Sortierspalten eingerichtet und nach welchen Prioritäten die Textsegmente schließlich geordnet werden. [10]

Nachdem schon verschiedentlich von Codes und Sortierkürzeln die Rede war, ist nun etwas genauer auf deren geeignete Struktur einzugehen, denn anders als ein QDA-Programm macht Word keinen prinzipiellen Unterschied zwischen den Primärdaten und den Codes und Sortierkürzeln. Damit Word aber keinen Verwechslungen aufsitzt, ist die Struktur der Sortierkürzel und Codes so festzulegen, dass sie sich zuverlässig von den Primärdaten abheben. Dies erreicht man mit der Verwendung von sogenannten *uncommon symbols* (FIELDING & LEE 1998), also von ungewöhnlichen Zeichenkombinationen aus Einzelbuchstaben, Zahlen und Sonderzeichen. Aus pragmatischen Gründen sollten möglichst einprägsame und kurze Codes und Sortierkürzel gewählt werden. Ein geeigneter Code für *Beteiligte und Betroffene* ist beispielsweise B&B: er ist für Word mit keiner Silbe in den Primärdaten verwechselbar, für den Kodierer leicht einprägsam und gleichzeitig kurz genug, um bei der Kodierung wenig Aufwand zu verursachen. [11]

3. Das breite Anwendungsspektrum des Such- und Ersetz-Befehls

Vom Word-Suchbefehl war schon verschiedentlich die Rede. Seine vielfältige Einsetzbarkeit für die technikunterstützte Analyse soll als nächstes etwas genauer dargestellt werden. [12]

Die einfachste Möglichkeit besteht darin, mit Hilfe des Suchbefehls jene Zitate und Textstellen in der Basisdatei abzurufen, welche schon während der Datensammlung als besonders aussagekräftig aufgefallen sind und deshalb in der Berichterstattung präsentiert werden sollen. Eine systematischere Möglichkeit besteht in der lexikalischen Suche, mit der nach bestimmten Buchstabenfolgen oder Worten gesucht wird. Sie ist generell zur Textexploration gut geeignet (KUCKARTZ 1999). Mit einer Suche nach der Zeichenfolge *Lehrer* zeigt Word am Bildschirm alle Textstellen an, welche die Worte *Lehrer, Lehrerin, Lehrerinnen, Lehrerverband* usw. enthalten. Hingegen wird der Begriff *Lehrperson* verpasst, weshalb man mit der kürzeren Zeichenfolge *Lehr* suchen könnte; damit werden nun aber nicht bloß die gewünschten Textstellen mit *Lehrperson* zusätzlich gefunden, sondern auch *lehren, Lehre, Lehrmittel* und so weiter. Dank der

Unterscheidung der Großschreibung bei den Sortierkürzeln lässt sich mit mancher Zeichenfolge noch gezielter suchen (*Bearbeiten / Suchen / Erweitern*). Eine kombinierte Suche mit mehreren Wörtern ist mit dem Such- und Ersetz-Befehl jedoch nicht möglich. [13]

Falls eine Liste der gefundenen Textstellen erstellt werden soll, so ist vor der Suche die Basisdatei in eine Tabelle umzuwandeln und eine entsprechende Sortierspalte hinzuzufügen, in welcher ein Sortierkürzel für jede gefundene Stelle manuell eingegeben wird. Nach der Sortierung ist dann leicht zu erkennen, in welchen Kontexten über Lehrer gesprochen wird. Manche QDA-Programme können solche Listen praktisch vollautomatisch erstellen, was insbesondere bei großen Datenbeständen komfortabler ist als die etwas umständliche Lösung mit Word. Der Mehraufwand bei Word entsteht v.a. durch das manuelle Kodieren der automatisch gefundenen Textstellen, wodurch die resultierende Liste dann wenigstens keine irrelevanten Hits enthält. [14]

Eine weitere Möglichkeit des Suchbefehls besteht im halbautomatischen induktiven Kodieren, welches zum computerunterstützten Kodieren von bestimmten Wortfamilien dient. Damit kann beispielsweise gezielt nach Kausalsätzen gesucht werden. Mit Hilfe der Suchworte *weil, da, wegen, darum, warum, deshalb* und *weshalb* zeigt Word der Reihe nach alle Textstellen am Bildschirm an, welche das gesuchte Wort enthalten. Mit einiger Wahrscheinlichkeit enthalten die so gefundenen Textsegmente Sätze mit Kausalaussagen, es werden sich darunter aber auch solche befinden, welche für die aktuelle Fragestellung irrelevant sind. Um für die weitere Analyse die Spreu vom Weizen zu trennen, erweitert man die relevanten Textstellen halbautomatisch mit einem Code. Dies geschieht mit dem Ersetz-Befehl unter *Bearbeiten / Suchen / Ersetzen*. Dabei ergänzt man in jedem relevanten Textsegment das Wort *weil* mit *weil Code* bzw. *wegen* mit *wegen Code* und so weiter. Word führt den Kodierer nach jedem so ausgeführten Ersetz-Befehl – bzw. nach jeder abgelehnten Ersetzung – zur jeweils nächsten Textstelle, die das Suchwort enthält. Dieses Vorgehen wird auf alle Suchworte der ausgewählten Wortfamilie angewandt. Am Schluss dieser gesamten Prozedur steht der Code in allen Textsegmenten der Basisdatei, die Word automatisch gefunden und der Kodierer als aussagekräftig beurteilt hat. Dank einem Suchbefehl mit dem Code lassen sich nun alle kodierten Kausalsätze der Reihe nach am Bildschirm anzeigen. [15]

Dieses halbautomatische induktive Kodieren ist sehr effizient, um einen raschen Überblick über ausgewählte Wortfamilien zu gewinnen. Es kann jedoch eine manuelle Inhaltsanalyse nicht ersetzen, sondern nur ergänzen, da jene Sätze, welche kein Suchwort aus der fraglichen Wortfamilie enthalten, nicht gefunden werden. Eine technische Einschränkung von Word besteht darin, dass die Fundstellen nur einzeln am Bildschirm angezeigt und nicht aufgelistet werden können. Wenn eine Liste aller kodierten Sätze erstellt und gedruckt werden soll, so ist analog wie bei den bereits erwähnten Sortiertabellen vorzugehen. [16]

Anders als solche induktive Codes werden a priori Codes nicht aus dem Datenmaterial herausentwickelt, sondern gemäß einem Codebuch den Daten manuell zugeordnet (HUBERMAN & MILES 1994). KUCKARTZ (1999) verwendet dafür den Begriff des intellektuellen Kodierens. Bei dieser Art des Kodierens schreibt man die entsprechenden Codes ans Ende des zu kodierenden Satzes oder Abschnitts. Alle a priori kodierten Textstellen können später dank dem Suchbefehl wiedergefunden und einzeln am Bildschirm angezeigt werden. Deren Auflistung bedingt wiederum die vorgängige Umwandlung der Basisdatei in eine Tabelle und den Eintrag geeigneter Codes in eine separate Sortierspalte. Das direkte Eingeben der Codes in die Primärdaten geht rascher als das formellere Vorgehen mit einer separaten Codespalte. [17]

Häufig besteht nach dem Kodieren die Notwendigkeit, die ursprünglich gewählten Kategorien feiner zu untergliedern, um eine detailliertere Analyse vorzunehmen. Dank dem Ersetz-Befehl kann man bestehende Codes und Sortierkürzel rasch anpassen. Wenn man beispielsweise zunächst für alle Konflikte den Code *Konlik*³ verwendet hat und nachträglich zwischen den Konfliktlinien 1 und 2 unterscheiden möchte, so lässt man sich alle *Konlik*-Textsegmente am Bildschirm anzeigen und ersetzt sie wahlweise mit den Subcodes *Konlik1* oder *Konlik2*. [18]

Gelegentlich will man zwei oder mehrere Codes miteinander fusionieren, v.a. wenn sie beim freien Kodieren entstanden sind (KUCKARTZ 1999). Dies lässt sich ebenfalls mit dem Ersetz-Befehl bewerkstelligen. Man ersetzt dann – analog dem eben erwähnten Beispiel – die Subcodes *Konlik1* bzw. *Konlik2* durch *Konlik*, wonach alle betroffenen Subkategorien wieder unter ein und derselben Kategorie figurieren. [19]

Überraschenderweise lässt sich der Ersetz-Befehl auch zum Auszählen von Häufigkeiten verwenden. Man braucht dazu lediglich den zu zählenden Code oder Begriff durch sich selbst zu ersetzen. Wenn beispielsweise die Häufigkeit des *CodeX* in der Basisdatei ausgezählt werden soll, so ersetzt man *CodeX* durch *CodeX*, was faktisch nichts ändert. Word führt diesen Ersetz-Befehl jedoch durch und zeigt die Anzahl der vorgenommenen "Ersetzungen" am Bildschirm an. Dank dieser Einsatzmöglichkeit des Ersetz-Befehls ist der Word-User zwar jeder manuellen Häufigkeitszählung weit überlegen, es lassen sich jedoch nicht verschiedene Codes oder Begriffe gleichzeitig auszählen oder gar Listen mit den verschiedenen Zählresultaten ausdrucken, wie das manche QDA-Programme können. Stattdessen ist für jede Zeichenkette eine separate Zählung auszulösen. Diese Nachteile im Vergleich zu einer Spezialsoftware fallen jedoch erst dann ins Gewicht, wenn die auszuzählenden Codes und Begriffe sehr zahlreich sind. [20]

3 Damit Word keinen Verwechslungen aufsitzt, ist die Struktur der Sortierkürzel und Codes so festzulegen, dass sie sich von den Primärdaten abhebt. Deshalb sollte man in diesem Beispiel nicht *Konflikt* als Code wählen, da diese Zeichenkette vermutlich auch in den Daten selber vorkommt. Beim leicht abgeänderten *Konlik* besteht kaum Verwechslungsgefahr zwischen Code und Primärdaten.

4. Die Kombination von verschiedenen Auswertungsoptionen

Um mit Word die gewünschten Resultate zu erzielen, ist für jedes Analyseproblem neu zu überlegen, welche Auswertungen sinnvoll sind und in welcher Reihenfolge dabei vorgegangen werden soll. Von Fall zu Fall können sich recht unterschiedliche Analysestrategien als sinnvoll erweisen. Als Beispiel stelle ich eine Studie vor, bei der ich als Einzelforscher verschiedene Auswertungsoptionen von Word miteinander kombinierte (NIDERÖST 2001). [21]

Als Primärdaten lagen für diese Fallstudie verschiedene Interviewtonbänder, die Akten einer parlamentarischen Kommission sowie eine größere Zahl von öffentlich zugänglichen Dokumenten und Medienberichten vor. Jene Textstellen der Primärdaten, welche für die Fragestellung besonders informativ waren, wurden in eine Word-Tabelle eingegeben, wobei zu jedem Segment ein Quellenvermerk erfasst wurde. Pro Tabellenspalte wurden die Daten zu je einem Akteur der Fallgeschichte erfasst. Mit der Transkription war also nebst der Datenerfassung auch ein Selektions- und Ordnungsprozess verbunden. In einer separaten Sortierspalte wurde zu jedem chronologisch bedeutsamen Textsegment das Datum des entsprechenden Ereignisses vermerkt, so dass jederzeit eine aktualisierte Version der Fallgeschichte erstellt werden konnte. Diese Organisation des Materials ermöglichte es zudem, während der laufend fortschreitenden Datensammlung stets den Überblick über den aktuellen Kenntnisstand zu behalten, da das Wiederfinden wichtiger Daten leicht war. In einer weiteren Tabellenspalte wurden Memos eingegeben. [22]

Nach der abgeschlossenen Datenerfassung wurde für die Analyse zwei Codelisten geschaffen und der Inhalt der Tabelle damit zweimal separat kodiert. Dabei wurden zwei unterschiedliche Kodierverfahren angewandt (HUBERMAN & MILES 1994). Es wurden mit beiden Listen dieselben Kategorien kodiert, jedoch mit formell unterschiedlichen Codes, damit Word zwischen den beiden Codelisten unterscheiden konnte. So konnten zwei unterschiedliche Kodierverfahren angewandt werden (siehe Abschnitt 3):

- a priori Codes: diese wurden den Primärdaten manuell zugeordnet,
- induktive Codes: diese wurden den Primärdaten aufgrund der Suchworte *weil, da, wegen, darum, warum, deshalb* und *weshalb* halbautomatisch zugeordnet. [23]

Für die Kategorie "Ermittlung aller Beteiligten und Betroffenen" wurde beispielsweise der a priori Code *B+B* bzw. der induktive Code *B&B* angewandt. Dem erhöhten Aufwand, der infolge des manuellen sowie halbautomatischen Kodierens der Basistabelle nötig war, standen mehr Objektivität sowie eine bessere Kontrolle der Reliabilität gegenüber, da auffällige prozentuale Differenzen zwischen den beiden Listen bei der späteren Auszählung der Codes sichtbar wurden. Ein analoges Verfahren ist weit verbreitet und besteht darin, die Primärdaten mit demselben Kodierverfahren und anhand einer einzigen Liste, jedoch durch zwei Personen unabhängig voneinander kodieren zu lassen und

anschließend deren Ergebnisse zu vergleichen. Dieses Vorgehen war jedoch bei einem Projekt mit einem einzigen Forscher nicht möglich; stattdessen übernahm Word beim halbautomatischen Kodieren sozusagen die Rolle des zweiten Kodierers. Schließlich wurden die Ergebnisse als vergleichende Auszählung der beiden Kodiersysteme sowie als dichte Beschreibung der einzelnen Kategorien präsentiert. [24]

5. Fazit

Als Programm, das für die Textverarbeitung geschaffen wurde, bietet Word nicht denselben Komfort und dieselbe Auswahl an Analysemöglichkeiten wie spezifische QDA-Programme. Die Anzeige von einzelnen Resultaten am Bildschirm muss häufig genügen, denn das Erstellen und Drucken von Ergebnislisten ist nur bei den tabellarischen Auswertungsoptionen möglich. Schließlich ist besonders sorgfältig auf die Festlegung von Kodier- und Verfahrenskonventionen zu achten, da Word nicht zwischen Primärdaten und Codes unterscheidet und nicht über vordefinierte Auswertungs- und Ausgaberroutinen verfügt wie QDA-Software. Hochspezifische Analyseschritte wie die kombinierte Suche nach mehreren Wörtern oder das Arbeiten mit Netzwerk-Kategoriensystemen sind mit Word nicht realisierbar, und auch sehr anspruchsvolle Auswertungsstrategien zur Identifikation von Mustern in den Daten oder hierarchisierte Retrieval-Funktionen bleiben für QDA-Programme wie WinMax oder ATLAS.ti reserviert. [25]

Trotz solchen Einschränkungen ist Word jedoch deutlich leistungsfähiger und vielseitiger, als man es einem Textverarbeitungsprogramm auf den ersten Blick zutrauen würde. Viele elementare Verfahren der technikerunterstützten Analyse sind machbar, wie beispielsweise einfache Code-and-Retrieve-Funktionen, Sortierungen, Auszählungen sowie das Aufteilen oder Fusionieren von Kategorien. Einfache Aufgaben, wie das Sortieren von Fragebogeninterviews oder Ordnen von Gruppengesprächsprotokollen, löst Word sehr effizient. Ein weiterer Vorteil besteht darin, dass das Importieren der Primärdaten in ein spezifisches QDA-Programm entfällt, womit auch alle Probleme mit Sonderzeichen, Unterstreichungen und anderen, nicht ASCII-fähigen Formatierungen wegfallen.⁴ Word ist aufgrund dieser Vor- und Nachteile geeignet für Projekte mit kleineren bis mittleren Datenmengen, für die lediglich einfachere Auswertungen erforderlich sind und deren Codelisten nicht sehr umfangreich sind. [26]

Die Hauptvorteile von Word liegen jedoch nicht auf der technischen Ebene, sondern sind in der Tatsache begründet, dass jede Forschungsarbeit bestimmten sozialen, personellen und finanziellen Restriktionen unterliegt. Als Standardsoftware ist Word auf den meisten Personalcomputern bereits installiert, womit man sich die Umstände und Kosten einer Neuanschaffung häufig ersparen

4 Die Entwickler von QDA-Software haben gerade in der jüngsten Vergangenheit an diesen Einlese-Problemen gearbeitet. In *Nvivo* können schon RTF-Dokumente importiert werden, und in der neuen *ATLAS.ti*-Version – die dieses Jahr noch auf den Markt kommen soll – kann man voraussichtlich alles einlesen, was in Word Pad dargestellt werden kann. Auch in *MAXqda*, dem neuen Programm der Entwickler von *winMAX*, können Texte in RTF-Format direkt importiert werden.

kann. Viele PC-Anwender sind mit Word sehr gut vertraut, so dass das Programm auch in der Praxis sofort benutzbar ist, da kaum Schulungsaufwand und Einarbeitungszeit erforderlich sind. Dank diesen sehr niedrigen Eintrittsschwellen stellt Word eine ideale Einstiegssoftware für all jene dar, die wenigstens einige grundlegende Optionen der technikerunterstützten Analyse nutzen wollen. [27]

Danksagung

Ich danke Stephanie SUMMERMATTER und Ursula HEGGLI für die Durchsicht des Manuskripts.

Literatur

[Fielding, Nigel G.](#) & Lee, Raymond M. (1998). *Computer Analysis and Qualitative Research*. London: Sage.

[Kelle, Udo](#) (Hrsg.) (1995). *Computer-Aided Qualitative Data Analysis. Theory, Methods and Practice*. London: Sage.

Kuckartz, Udo (1999). *Computergestützte Analyse qualitativer Daten. Eine Einführung in Methoden und Arbeitstechniken*. Opladen: Westdeutscher Verlag.

Nideröst, Bruno (2001). *Eine erfolgreiche Politikevaluation. Die Evaluation der Eidgenössischen Volkszählung und die Gründe für deren Umsetzung*. Universität Bern, Institut für Politikwissenschaft.

Miles, Matthew B. & Huberman, A. Michael (1994). *Qualitative Data Analysis*. London: Sage.

Zum Autor

Bruno NIDERÖST ist Politikwissenschaftler mit den Arbeitsschwerpunkten Evaluation, Politikanalyse und Verfassungsgeschichte. Er forscht am Institut für Politikwissenschaft der Universität Zürich und dem Institut für Öffentliches Recht der Universität Bern.

Kontakt:

Bruno Nideröst
Seestrasse 48
6424 Lauerz, Schweiz

E-Mail: bnideroest@dplanet.ch

Zitation

Nideröst, Bruno (2002). Die technikerunterstützte Analyse von qualitativen Daten mit Word [27 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 22, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0202225>.

Revised 2/2007