

149. Mergenthaler E, Kächele H (1994) Die Ulmer Textbank.
Psychother psychol Med 44: 29-35

Die Ulmer Textbank

Erhard Mergenthaler; Horst Kächele

Sektion Informatik in der Psychotherapie an der Abteilung Psychotherapie,
Universität Ulm

Key words: Process research, textbank, computer-technology

Dr. Erhard Mergenthaler
Sektion Informatik in der Psychotherapie
Universität Ulm - Klinikum
Am Hochsträß 8
D-7900 Ulm

Abstract

The Ulm Textbank

Since 1968 one of our major research efforts consists in establishing a methodology for performing psychoanalytic process research. Within this frame tape-recording of psychoanalytic long-term treatments constituted an essential methodical step inevitably leading to the production of a large collection of verbatim transcripts. We gradually and inadvertently realized the need for a major computerized databank to assist our own research. With support of the German Research Foundation we started in 1980 with the development of the Ulm Textbank Management System. While realizing the system it became obvious that such a databank would serve as well other researchers involved in process research when analyzing verbatim material. The final shape of the systems thus was strongly influenced by the orientation towards a variety of users and methodological approaches. Meanwhile this task is completed and the ULM TEXTBANK, as it is known, is available as a new unique tool for psychotherapy research.

Zusammenfassung

Die Ulmer Textbank entwickelte sich als Folge der 1968 in Ulm begonnenen psychoanalytischen Prozessforschung anhand von Tonbandaufzeichnungen. Hieraus entstand die Konzeption einer Datenbank, die sowohl archivarische als auch datenanalytische Funktionen würde übernehmen können. Dieses Vorhaben wurde von 1980 bis 1988 mit Unterstützung der DFG verwirklicht. Dieser Beitrag gibt einen Überblick über Entstehung und Leistungen der Ulmer Textbank

1. Ziele des Textbankprojektes

Umfangreiche verbatim-transkribierte Gesprächsprotokolle haben sich als eine wichtige Datenquelle in der psychotherapeutischen Forschung etabliert, wie dies bereits Luborsky und Spence (1971) gefordert haben. Aus heutiger Sicht zeigt sich deutlich, daß im Hinblick auf vielfältige Erwartungen es schon lange anstand, für den Anwendungsbereich Psychotherapie geeignete und benutzerfreundliche Methoden zur Handhabung eines Textkorpus zu entwickeln. Darüber hinaus wird auch offenbar, wie wichtig es war, aussagekräftige Methoden zur Beschreibung solcher Texte zu entwickeln oder aus der linguistischen Datenverarbeitung zu übernehmen. Zur Lösung der anfallenden Probleme wurde vor zwanzig Jahren in Ulm ein interdisziplinärer Ansatz gewählt, der die psychotherapie-bezogenen Fragestellungen mit wissenschaftlichen Methoden der Informatik und Linguistik verbindet

1.1 Geschichtlicher Abriß

Eine der Hauptforschungsanstrengungen an der Abteilung Psychotherapie der Universität Ulm richtet sich seit 1968 auf die Entwicklung einer Methodologie für psychoanalytische Prozeßforschung. Innerhalb dieses Rahmens bilden Ton- und Videoaufnahmen von psychoanalytischen Langzeitbehandlungen einen wesentlichen methodologischen Schritt, der unvermeidlich zur Produktion einer großen Sammlung von Verbatim-Transkripten führte. Erst allmählich realisierten wir die Notwendigkeit, eine größere computergestützte Datenbank für unsere Forschung zu entwickeln. So begann mit dem Sonderforschungsbereich 129 auch die Entwicklung des Ulmer-Textbank-Verwaltungssystems. Während dieser Entwicklungszeit zeigte sich zudem, daß eine solche Datenbank auch anderen an der Prozeßforschung interessierten Wissenschaftlern helfen würde, sprachliches Material zu analysieren. Die letztendliche Ausgestaltung des Systems war daher stark von der Orientierung an einer vielschichtigen Benutzerschaft mit sehr unterschiedlichen methodologischen Ansätzen geprägt. Mit dem Abschluß des SFB 129 konnte auch diese Aufgabe beendet werden. Das Projekt steht heute als ULMER TEXTBANK für die Psychotherapieforschung zur Verfügung.

1.2 Allgemeine Ziele

Eines der Hauptziele bei der Entwicklung der Textbank war es, vielen verschiedenen Forschern Sprachmaterial von psychotherapeutischen Sitzungen aber auch aus benachbarten Feldern verfügbar zu machen, um Zeit und Geld für solche Forschungsvorhaben zu sparen, die mit bereits verfügbarem Material durchgeführt werden können (Bibliotheks- und Archivfunktion). Ein weiteres Ziel war, nicht nur Verbatimmaterial zur Verfügung zu stellen, sondern auch Zugang zu computergestützten Textanalysen zu schaffen für all die Wissenschaftler, die über keine eigenen entsprechenden Computerressourcen verfügen. Ein drittes Ziel schließlich bestand darin, Ergebnisse, die in vorausgegangenen Analysen an diesen Daten gewonnen wurden, miteinander zu verbinden und so ein Wiederauffinden von Texten auf der Basis der bereits vorliegenden Ergebnisse zu ermöglichen. Entsprechend wurde der Entwurf des Textbankverwaltungssystems an folgenden Aufgaben ausgerichtet:

- a) Erfassen und Aufbereiten von Texten unter vielfältigen Gesichtspunkten.
- b) Verwalten beliebig vieler Texteinheiten auf dem Hintergrundspeicher einer Rechananlage.
- c) Verwalten von beliebig vielen Informationen über die Texteinheiten und deren Autoren sowie über die an ihnen durchgeführten Textanalysen.
- d) Verwalten einer offenen Menge von Methoden zur Verarbeitung und Analyse der gespeicherten Texteinheiten.
- e) Unterstützung von Schnittstellen zu Statistik- und sonstigen Anwenderpaketen.
- f) Unterstützung einer einfachen interaktiven Benutzerschnittstelle bei der Inanspruchnahme oder Durchführung der unter a) bis e) genannten Aufgaben.

Das Textbankverwaltungssystem ist damit ein Informationssystem, das Texte und Informationen über Texte verwalten kann und Verfahren der linguistischen Datenverarbeitung sowie der Textverarbeitung zur Texterschließung integriert. Es ist gekennzeichnet durch eine homogene Benut-

zerschnittstelle, die Aufnahme, Verarbeitung, Ausgabe und Analyse der Texteinheiten im Dialog unterstützt.

Die in der ULMER TEXTBANK gespeicherten Dokumente sind vornehmlich eine offene Sammlung von Texten. Das Hauptmerkmal solcher Datensammlungen ist, daß sie kontinuierlich erweitert werden können. Das Maß an Vollständigkeit einer Datenbank beeinflußt aber auch die Strategien bei der Handhabung von Forschungsergebnissen zu diesen Texten. Zwei Ansätze können unterschieden werden: Beim ersten werden alle verfügbaren Daten zusammen mit dem Text selbst gespeichert, beim anderen werden die Analysen je nach Bedarf erneut durchgeführt.

1.3 Weitere Forschungsziele

Das Textbankprojekt repräsentiert die Verwirklichung von Informatikwerkzeugen in der Psychotherapieforschung. Besonderes Interesse gilt in diesem Zusammenhang der Akzeptanz und Performanz dieses recht neuen Ansatzes und dem Sammeln von Erfahrungen. Während der Phase des Ansammelns von Text mußte außerdem das Feld mit einem neuen Faktum, nämlich dem gemeinsamen Nutzen von Primärdaten vertraut gemacht werden. So war es einer der lohnendsten Aspekte des Projektes festzustellen, daß eine rasch anwachsende Zahl von Kollegen unsere Zielsetzung verstand sich ihr anschloß, und in großzügiger Weise zum Gelingen beitrug, indem sie ihre Datenquellen zur Verfügung stellten¹.

2. Methoden

2.1 Klientel und Stichproben

Der optimale Einsatz eines Textbankverwaltungssystems in der Psychotherapieforschung erfordert, daß die zu verwaltende Textbasis es erlaubt, mögliche, im einzelnen schwer vorherzusehenden Forschungsfragestellungen bearbeiten zu können. Es ist daher besonders wichtig, daß individuelle Textsammlungen als Untereinheiten der Textbank zusammengestellt werden können. In diesem Zusammenhang haben sich zwei Arbeitsschwerpunkte der ULMER TEXTBANK herauskristallisiert, die zugleich

¹Unser Dank gilt den vielen Kolleginnen und Kollegen, die uns ihr Vertrauen und ihre Kooperation geschenkt haben.

zwei unterschiedlichen Forschungsansätzen entsprechen: Längsschnittstudien und Querschnittstudien.

Die Längsschnittstudien konzentrieren sich auf Texte aus psychotherapeutischen und psychoanalytischen Behandlungen. Ihr Ziel ist die Erforschung der Veränderung durch den therapeutischen Prozeß. Wegen der Vielzahl der Stunden, insbesondere bei hochfrequenten Psychoanalysen können Transkripte nur von einer geringen Anzahl verschiedener Behandlungen bearbeitet werden. Es stehen deshalb Auswertungen zu vielseitigen Veränderungen in Einzelfällen im Vordergrund.

Natürlich gibt es auch Fragen, die über die Veränderung bei einzelnen Patienten oder Therapeuten hinausgehen. Diese werden in Querschnittsuntersuchungen an Erstinterviewtexten bearbeitet. Dies bedeutet, daß viele verschiedene Patienten mit jeweils nur einem Interview untersucht werden können, wodurch es möglich wird, den Einfluß von Variablen, wie beispielsweise Geschlecht oder Diagnose zu beobachten. Daneben werden gesonderte Textsammlungen geführt, die für spezielle Untersuchungen, wie etwa der Erforschung von Balint-Gruppen, des Sprachverhaltens während Visitengesprächen oder des sprachlichen Austauschs in Familientherapien benötigt werden.

Texte, die den Hauptzielen der ULMER TEXTBANK entsprechen, werden systematisch ergänzt. Das Archiv enthält mittlerweile neben mehrerer vollständig verfügbarer Kurztherapien auch umfangreiche Stichproben zu vier psychoanalytischen Behandlungen. Das Erstinterviewkorpus besteht aus mehreren hundert verschiedenen Interviews und ist hinsichtlich Geschlecht der Patienten bzw. Therapeuten und hinsichtlich der diagnostischen Unterscheidung Neurose bzw. psychosomatische Störung ausgeglichen. Dieses Textkorpus wird nach wie vor vergrößert, insbesondere im Hinblick auf die Patienten-Variablen Geschlecht, Diagnose, sozialer Status und Alter, der Therapeuten-Variablen Erfahrung und außerdem hinsichtlich der Therapieart.

Die Art von Texten, wie sie in der Textbank vorzufinden sind, bestimmte auch die Ziele, Fragestellungen und wissenschaftlichen Interessen der anderen unterstützenden Einrichtungen. Für die Abteilung Psychotherapie der Universität Ulm bedeutet dies sowohl die Schaffung einer empirischen

Basis für die Forschung in der Psychotherapie als auch die Unterstützung in der Lehre. Das letztere umfaßt Demonstrationsmaterial für die Ausbildung von Medizinstudenten und den Gebrauch von Verbatimprotokollen bei der klinischen Ausbildung und Supervision von Ärzten und Psychologen.

Zwei Drittel des Materials der ULMER TEXTBANK kommt aus Ulm selbst. Die übrigen Texte wurden als Ergebnis wissenschaftlicher Kontakte und gemeinsamer Forschungsprojekte mit Einrichtungen außerhalb von Ulm gewonnen. In den meisten Fällen war die Überlassung der Texte gebunden an die Inanspruchnahme von Dienstleistungen über das Textbanksystem. Während diese vornehmlich aus dem engeren Umfeld der Psychotherapie stammten, gehörten die übrigen Nutzer vorwiegend zu den Linguisten und Sozialwissenschaftlern. Derzeit bestehen Kontakte zu ungefähr dreißig Instituten in Deutschland, vier in den Vereinigten Staaten, zwei in Schweden, zwei in der Schweiz und einem in Österreich. Insgesamt umfassen die elektronisch gespeicherten Texte ein Vokabular von 180.000 verschiedenen deutschen Wörtern mit einer Gesamtauftrittshäufigkeit von mehr als 10 Millionen Wörtern (Textumfang).

Fragen hinsichtlich der Repräsentativität der ULMER TEXTBANK orientieren sich an den Forschungszielen. Allerdings gibt es praktische Grenzen, wie sie etwa durch die große Anzahl von Behandlungsstunden in Psychoanalysen bedingt sind. Bei der Auswahl der Einzelfälle für die Textbank waren jedoch nicht nur praktische Gesichtspunkte von Bedeutung. In der Reihenfolge ihrer Wichtigkeit sind dies: Die zahlenmäßige Ausgewogenheit der verschiedenen Therapeuten, der diagnostischen Kategorien, Behandlungen mit einer Gesamtstundenzahl zwischen 300 und 500 und dem Therapieerfolg. Andere Auswahlkriterien, wie sie für statistische Auswertungen von Bedeutung wären, können aufgrund der geringen zur Verfügung stehenden Anzahl von Fällen nicht berücksichtigt werden. Demzufolge kann das Psychoanalysekörpus in Ulm nur im Hinblick auf spezielle Fragestellungen als repräsentativ betrachtet werden. Die Tabelle im Anhang gibt eine Bestandsübersicht zum Ende des Jahres 1992.

2.2 Instrumente

Ausgehend von einer semiotischen Sicht der Sprache, wie sie sich auf Peirce, dem Begründer der Semiotik und ihren Weiterentwicklungen durch Morris zurückführen läßt, wird Sprache als ein System von Symbolen verstanden, dessen Struktur bestimmt wird durch Regeln über die Beziehung zwischen Form und Inhalt.

Entsprechend ist es möglich, zwischen formalen, grammatischen und inhaltlichen Textmaßen zu unterscheiden. Jede dieser Meßmethoden kann weiter unterschieden werden im Hinblick auf den Beitrag eines einzelnen Sprechers oder auf den Text als Ganzes, als Dialog. Es kann daher von monadischen oder dyadischen Meßwerten gesprochen werden. Außerdem kann nach der Art der Meßwerte unterschieden werden. Wohl bekannt sind die einfachen Maße der Auftretenshäufigkeit, die wiederum die Basis für Verhältniszahlen und Verteilungen bilden. Weiterhin ist zu beachten, daß entsprechend der Unterscheidung, wie sie hier gemacht wurde, einige der Ansätze für formale und grammatische Maße inhaltliches Wissen voraussetzen, wie etwa die denotative Bedeutung eines Wortes. Im Unterschied zu den inhaltlichen Maßen kommt dieses Wissen jedoch nicht aus dem Forschungsfeld selbst, nämlich der Psychoanalyse, sondern aus den methodologischen Bereichen, wie der Linguistik oder Informatik.

Die formalen Maße können im allgemeinen auf sehr einfache Art und Weise ermittelt werden. Bei computergestützten Ansätzen ist lediglich die Fähigkeit zur Segmentierung von Symbolfolgen (Buchstaben, Ziffern und Sonderzeichen) zu Wörtern und Satzzeichen erforderlich. Der Programmieraufwand ist verhältnismäßig gering, Umkodieren oder Präkodieren ist nicht notwendig. Eine Aufstellung solcher formalen Maße, wie sie auch über die Methodenbankkomponente der ULMER TEXTBANK zur Verfügung stehen, umfaßt: Textumfang (Tokens), Vokabular (Types), Typ-Token-Ratio, Redundanz und Sprecherwechsel bei Familien- und Gruppengesprächen.

Das einfachste und elementarste formale Maß ist die Anzahl der vom Therapeuten oder Patienten gesprochenen Wörter. Kächele (1983) fand heraus, daß in einer erfolgreichen psychoanalytischen Behandlung in einer Stichprobe von über 130 Stunden keine Korrelation zwischen der Anzahl der vom Patienten und Analytiker geäußerten Wörter bestand (Pat. Amalie

X²). In einer zum Zeitpunkt der Studie noch nicht erfolgreich abgeschlossenen Behandlung durch denselben Analytiker waren diese Wortzählungen signifikant positiv korreliert (+.30 bei N=110; Pat. Christian Y). O'Dell und Winder (1975) verwendeten ebenfalls den Textumfang als ein Maß für die Aktivität des Therapeuten, um verschiedene therapeutische Techniken zu unterscheiden. Sie benennen 7 % als Anteil des Therapeuten in analytischer Therapie und 31 % in eklektischer Psychotherapie. Zimmer und Cowles (1972) führten eine Studie durch mit einem Patienten, der bei drei Therapeuten verschiedener theoretischer Orientierung gesehen wurde. Auch sie stellten signifikante Unterschiede fest. Gestützt auf dieselben Daten zeigte Pepinsky (1979), daß die Art und Ausprägung der Aktivität des Therapeuten den Patientin dahingehend beeinflußt, sich sprachlich in einer ähnlichen Weise wie der Therapeut zu verhalten.

Die Redundanz ist ein Textmaß, das aus der Informationstheorie stammt. Spence (1968) hat einige grundlegende Überlegungen zur psychodynamischen Redundanz angestellt, ohne dies jedoch selbst empirisch überprüfen zu können. Außerdem formulierte er eine Reihe von Hypothesen über den Verlauf der Redundanz während einer psychoanalytischen Behandlung. Kächele und Mergenthaler (1984) haben eine dieser Hypothesen, nämlich, daß die sprachliche Redundanz des Patienten (häufigeres Wiederholen von Wörtern) während einer Behandlung zunähme, bestätigt. Die Werte des Therapeuten dagegen blieben konstant.

Die grammatischen Maße verlangen vom Wissenschaftler, daß er linguistisches Wissen über die zu untersuchende Sprache, z. B. über die Grammatik des Deutschen, besitzt. Der Programmier- und Kodieraufwand für computergestützte Prozeduren ist bereits beträchtlich. Mehr noch, viele Fragen können bis heute noch nicht vollständig automatisch bearbeitet werden. Ein Beispiel ist die Lemmatisierung, also das automatische Zurückführen einer flektierten Wortform auf ihre Grundform, das heute je nach Textart einen Wirkungsgrad zwischen 50 und 95 % aufweist. Das psychotherapeutische Gespräch, eine Sprachform mit vielen syntaktischen Abweichungen (beispielsweise unvollständige Wörter und Sätze), ist typisch für gesprochene und spontane Sprache und rangiert damit im unteren

²Die Namen beziehen sich auf die klinischen Beschreibungen der Patienten, die in Thomä & Kächele (1988) eingeführt wurden.

Bereich. Entsprechend gibt es auch kaum computergestützte Analysen von psychotherapeutischen Texten, die sich auf grammatische Maße stützen. Eine über die Methodenbank-Komponente der ULMER TEXTBANK verfügbare Auswahl solcher Maße ist: Wortartenverteilung, Minderung und Steigerung und Interjektionen.

Die Wortartenverteilung wurde von Lorenz und Cobb (1954) dazu herangezogen, um zwischen Patienten verschiedener psychotischer Erkrankungen zu differenzieren. Ein Ergebnis ihrer Untersuchungen war, daß Neurotiker mehr Verben aber weniger Konjunktionen im Vergleich zu einer Normalpopulation benutzten. Allerdings müssen hier, wie Eisenmann (1973) für Konjunktionen zeigen konnte, weitere Variablen beachtet werden. So bestimmt sich der Gebrauch der Konjunktionen vornehmlich durch geografische Einflüsse, durch Geschlecht, durch Alter und schließlich durch die soziale Schichtzugehörigkeit.

Der Zusammenhang zwischen der Wahl einer Wortart und der zugehörigen semantischen Klasse wurde von Busemann bereits 1925 in einer Untersuchung der Kindersprache gezeigt. Er sprach von einem "aktiven" und "qualitativen" Stil in Bezug zu Verben bzw. Adjektiven. Er zeigte weiterhin, daß diese Stilunterschiede nur geringfügig vom Thema über das geprochene Wort abhängig sind und eher zum Bereich der persönlichen Variablen gerechnet werden müssen. In einem computergestützten Ansatz zeigten Mergenthaler und Kächele (1985) an einem psychoanalytischen Gespräch, daß die Wahl der Wortart definitiv vom zu berichtenden Inhalt abhängt. Allerdings schließt diese mikroanalytische Sicht die Möglichkeit nicht aus, daß von einer Makroebene aus gesehen, Persönlichkeitsvariablen, wie sie von Busemann beschrieben wurden, einen Einfluß haben können.

Der Verb-Adjektiv-Quotient, wie er von Boder (1940) analog zu Busemanns Aktionsquotienten eingeführt wurde, wurde von Wirtz & Kächele (1983) auf eine Stichprobe von Erstinterviews von drei verschiedenen Therapeuten angewendet. Sie zogen den Schluß, daß dieser Quotient sowohl ein differentielles Maß für Sprechstile des Therapeuten als auch hinsichtlich Geschlecht und Diagnose sein kann.

Die Bedeutung von Personalpronomina für die Strukturierung von Selbst- und Objektbeziehungen in der Sprache wurde von Cierpka et al. (1980, 1983) an Texten einer psychoanalytischen Gruppentherapie gezeigt; Schaumburg (1980) analysierte den Gebrauch an den vier Musterfällen der Ulmer Korpus.

Angstthemen oder Primär-/Sekundärprozeßanteile können als Beispiele für inhaltliche Maße herangezogen werden. Sie erfordern zusätzliches detailliertes Expertenwissen hinsichtlich der verwendeten Theorien im Anwendungsbereich. Computergestützte Verfahren können hier nur annäherungsweise Ergebnisse liefern und sind auf engumrissene Konstrukte begrenzt. Neuere Arbeitsfelder der Informatik, insbesondere der Bereich der künstlichen Intelligenz könnten hier in absehbarer Zeit einen Durchbruch ermöglichen, indem wissensbasierte Ansätze zusehends zum Tragen kommen. Hier kommen insbesondere regelorientierte Ansätze in Betracht (Clippinger 1977, Teller und Dahl 1981).

Die bisher bedeutendste inhaltlich orientierte Methode geht auf Gottschalk und Gleser (1969, 1974) zurück, die die im Bereich der Psychotherapie am weitesten verbreiteten Skalen entwickelt haben. Eine Überarbeitung für das Deutsche wurde von Koch und Schöfer (1986) vorgenommen. Ein Vergleich dieses Verfahrens mit computergestützten Ansätzen gaben Grünzig & Mergenthaler (1986) sowie Lolas et al (1982), die am selben Textmaterial Unterschiede der intellektuellen und computergestützten Inhaltsanalyse herausgearbeitet haben. In einer wegweisenden Studie hat Dahl (1972) den Negativkurs einer 363 Stunden umfassenden und zweieinhalb Jahre dauernden nicht-erfolgreichen Psychoanalyse überzeugend an einer Stichprobe von 25 Stunden kategorisiert, wobei er zehn extreme Arbeitsstunden, zehn extreme Widerstandsstunden und fünf Stunden zwischen diesen Extremen herausgestellt hat. Unter Bezugnahme auf einzelne Wörter aus dem Harvard-III-Diktionär konnte er Wortcluster herausstellen, die ödipale und andere unbewußte Konflikte klar repräsentieren (Dahl 1974). Diese Studie überzeugte uns, diesen Ansatz ebenfalls zu übernehmen.

Gestützt auf eine deutsche Adaption des Harvard-III-Psychosozialen-Diktionärs konnte Kächele (1976, 1988) demonstrieren, daß lineare Kombinationen von Kategorien hochgradig klinische Konzepte wie positive und

negative Übertragungskonstellationen in Verbindung mit ausgewählten Angstthemen vorhersagen können. Seine Ergebnisse stützten sich auf eine Einzelfallstudie (am Patient Christian Y) mit einer Stichprobe von insgesamt 55 Stunden und Korrelationen, die zwischen .77 und .91 lagen.

Reynes et al. (1984) benutzten das Regressive-Imagery-Dictionary (RID), um die bereits o.e. zehn Widerstands- bzw. Arbeitsstunden zu vergleichen. Die Arbeitsstunden waren charakterisiert durch eine Zunahme der Kategorien bezüglich Primärprozeß, die Widerstandsstunden waren gekennzeichnet durch einen höheren Wert an sekundär-prozeßbezogenen Kategorien. Dies stimmt mit Freuds früher Attribution von Abwehrfunktionen zum Sekundärprozeß überein (Bucci 1988).

Große zusammenhängende Textsegmente aber auch ausgewählte Abschnitte aus Behandlungsprotokollen können also mit Hilfe computergestützter Textanalyse als Werkzeug der psychoanalytischen Prozeßforschung untersucht werden (Kächele und Mergenthaler 1984). Die Fortführung dieser Ansätze bedarf jedoch, daß die vorhandenen Methoden intensiv weiterentwickelt, Grundlagenforschung fortgeführt und Techniken benachbarter wissenschaftlicher Disziplinen, wie der Informatik und Linguistik, mit einbezogen werden.

2.3 Randbedingungen

Bei der Aufnahme eines Textes in die Textbank werden alle Eigennamen, Landschaftsbezeichnungen und sonstigen persönlichen Merkmale mit Hilfe kryptografischer Prozeduren verschlüsselt oder durch Pseudonyme ersetzt. Während die Texte, die auf diese Art und Weise faktisch anonym sind, im Abteilungsrechner (SUN Server mit lokalem Macintoshnetz) verarbeitet werden können, sind die Schlüsseldaten, also alle persönlichen Angaben, auf Arbeitsplatzrechnern, die ausschließlich der ULMER TEXTBANK zur Verfügung stehen, davon getrennt gespeichert. Diese getrennte Datenhaltung sowie extensive Kontrollmechanismen schützen die ULMER TEXTBANK weitgehend gegen Mißbrauch. Das an der Textbank beschäftigte Personal ist auf den Datenschutz verpflichtet.

Für weitere technische Details, die die Struktur, den Datenbestand oder den Gebrauch der ULMER TEXTBANK betreffen, wird auf Mergenthaler (1986) verwiesen.

3.

3.1 Perspektiven bisheriger Forschungsergebnisse

Der methodologische Ansatz des Projektes erlaubte bis heute, eine Vielfalt von Analysen und Studien mit Hilfe der ULMER TEXTBANK und der von ihr angebotenen Methodik durchzuführen. Die Dienste wurden in Anspruch genommen von Wissenschaftlern, die sich mit Gruppendynamik, Familieninteraktion, Einzelpsychotherapie, Prozeßforschung, Balintgruppen, genetischer Beratung, soziologischen Interviews und psychiatrischen Interviews befaßten. Unsere eigene Arbeit hatte den Schwerpunkt in der psychoanalytischen Prozeßforschung und stand unter dem Aspekt der Entwicklung von Forschungsinstrumenten. Der übernationale Aspekt des Textbankprojekts kann am Beispiel einer Studie verdeutlicht werden, die wir an Luborskys 10 "Improver" und "Non-Improver"-Fällen (Philadelphia) durchgeführt haben, indem wir das Vokabular dieser Stunden unter dem Aspekt der "hilfreichen Beziehung" untersuchten (Hölzer et. al., in Vorb.)

3.2 Laufende Forschungsarbeiten

Das von der Deutschen Forschungsgemeinschaft im Rahmen des SFB 129 von 1980 - 1988 großzügig geförderte Projekt der ULMER TEXTBANK endete nicht mit dem Ablauf der Förderperiode, sondern wurde als "Sektion Informatik in der Psychotherapie" an der Abteilung Psychotherapie in die Grundausrüstung des Klinikums der Universität Ulm übernommen. Damit wurde eine Fortsetzung der Forschungsarbeiten sichergestellt, von denen einige im Folgenden skizziert werden sollen.

a.. Veränderung des Vokabulars

Eine Studie zu Veränderungen des Vokabulars im Laufe einer psychoanalytischen Kurztherapie ("Der Student"), insbesondere unter dem Aspekt der gefühlsbetonten Wörter, ist abgeschlossen. Hierfür wurde ein Diktionär zur Messung von Emotionen (UAD), gestützt auf Dahls Emotionstheorie (Dahl, 1988), entwickelt, das spezifische Änderungen in der Verbalisierung von Emotionen, wie sie durch die klinische Arbeit bewirkt werden, beschreibt (Hölzer et al., 1990). Das ADU wird derzeit auch zur Identifizierung von "frames" herangezogen (Hölzer et al. 1993

b. Untersuchung der Wortartenverteilung:

Substantive, Verben und Adjektive sind die am häufigsten verwendeten Wortarten. Jede von Ihnen repräsentiert bedeutende, zugrundeliegende Konzepte wie Zustände, Aktivitäten und Eigenschaften. Es ist bereits klar, daß die Verteilung dieser Variablen über die Zeit sich ändert und zwar entsprechend dem Verlauf des therapeutischen Prozesses und der damit einhergehenden Emotionen. Im Rahmen einer Pilotstudie wurde dieses Maß auch auf die Antworten von Klienten auf einen projektiven psychologischen Test angewendet. Zwei Gruppen von jeweils 30 als neurotisch- bzw. borderline-klassifizierten Patienten konnten klar differenziert werden. Der wichtigste Befund war ein signifikant häufigerer Gebrauch von Verben bei den Neurotikern (Mergenthaler & Pokorny, 1989).

c. Sprache und Physiologie:

Die Ergebnisse einer Pilotstudie, welche die Herzrate mit verschiedenen Textmaßen in Verbindung brachte, geben Anlaß, diese Fragestellungen weiter zu verfolgen. Die Ergebnisse wurden an konfliktbehafteten und verdrängten Themen interpretiert, wie sie auch über andere methodologische Ansätze herausgearbeitet wurden. Es gibt Hinweise auf eine signifikante Konvergenz von computergestützten Textanalysen und anderen Ansätzen. Diese Studie wird in Zusammenarbeit mit dem Programm zur "Untersuchung von bewußten und unbewußten mentalen Prozessen" (PCUMP, M. Horowitz) an der Universität von Californien, San Francisco durchgeführt (Mergenthaler 1992, Horowitz et al. in press).

d. Die Analyse von Sprecherabfolgen in Familien- und Gruppentherapien:

Es wurden mathematische Modelle und Prozeduren für die grafische Darstellung von Sprecherabfolgen entwickelt und auf eine Reihe von Familientherapien angewendet. Die Ergebnisse zeigen klare Unterschiede zwischen einzelnen Familien. Im Hinblick auf diagnostische Kriterien jedoch konnte diese Technik keine systematischen Erkenntnisse bis heute zeigen (Brunner 1985).

e. Computergestützte Erfassung von Bucci's referentieller Aktivität

Erste Ansätze an Therapien aus dem PCUMP-Projekt (Horowitz et al.) wirken erfolgversprechend. Das Projekt soll, sowohl fürs Englische als auch fürs Deutsche, Verfahren bereitstellen, um dieses Ausmaß der Bezugnahme auf das visuelle Gedächtnis im Sprachgebrauch messen zu können.

5. Zukunftspläne

5.1 Neue Forschungsfelder und -richtungen

Die computergestützte Textanalyse hat sich als ein wertvolles Werkzeug in der psychoanalytischen, allgemein der psychotherapeutischen Therapieforschung erwiesen (Kächele und Mergenthaler 1983, 1984). Auffallend ist, wie wenig Kooperationen sich bislang mit psychiatrischen Forschungsansätzen ergeben haben. Nur in einer Studie konnten wir Hilfestellung bei der Untersuchung von verbaler Interaktion bei psychiatrischen Patienten in einer Poliklinik leisten (Scheibe 1991). So dürfte es eine Aufgabe sein, verstärkt diese Kontakte zu suchen und die Sichtbarkeit und Zugänglichkeit der Ulmer Textbank zu erhöhen.

Weitere methodische Fortschritte sind nur durch die Überwindung der Schwachstellen der gegenwärtigen Untersuchungstechniken zu erreichen. Dies beginnt mit dem Prozess der Datengewinnung, der nach wie vor mit dem zeitaufwendigen Transkribieren verknüpft ist, jedoch mittlerweile durch die Entwicklung von Standards bereits wesentlich effizienter und zuverlässiger durchgeführt werden kann (Mergenthaler, 1986; Mergenthaler und Stinson, 1992). Die weiteren Schritte der qualitativen und quantitativen Texterschließung werden in absehbarer Zeit durch sog. Hyper-Text Ansätze wesentlich erweitert, indem in äußerst komfortabler Form Werkzeuge zur Archivierung, zum Wiederfinden (retrieval), Analysieren und Attribuieren von Texten zur Verfügung gestellt werden können

Die Ulmer Textbank begann als "big science" Unternehmen in der Großrechnerwelt. Die zwischenzeitlich erfolgte Entwicklung der PC-Welten läßt erwarten, daß Textanalysesysteme zunehmend in den alltäglichen Raum des Wissenschaftlers eindringen und sich für umschriebene Analy-

sen anbieten. Die zentralistische Perspektive bleibt aber erhalten, da die Notwendigkeit weiterer Entwicklungsarbeiten mehr denn je besteht.

6. Der Bezug zu anderen Forschungsprogrammen

Die Dienste der ULMER TEXTBANK stehen bei nur geringen Kosten auch anderen wissenschaftlichen Einrichtungen zur Verfügung. Gebühren werden insbesondere für arbeitsintensive Aufgaben, wie beispielsweise die Transkription von tonbandaufgenommenen Gesprächen sowie für Material erhoben. Andererseits wird erwartet, daß Texte, die auf diese Weise ihren Weg in die Textbank finden, auch für andere Wissenschaftler zukünftig zugänglich sind. Im Hinblick auf Material, das von der Textbank ausgeliehen wird, sollte eine Kopie des Berichtes oder der Publikation zurückfließen. So kann zusätzlich zu den Texten ein wachsender Bestand an Wissen über die Texte von unterschiedlichen wissenschaftlichen Disziplinen gespeichert und anderen zur Verfügung gestellt werden. Die ULMER TEXTBANK ist offen für alle, die ihre Texte dort speichern wollen. Allein die Möglichkeit routinemäßiger oder aber auch maßgeschneiderter Textanalysen, die einfache Art der Textverwaltung oder die Möglichkeiten der vielfältigen Druckausgaben sollten Anreiz genug sein, diese Dienste in Anspruch zu nehmen.

Literatur

- Boder DP (1940) The adjective-verb-quotient: a contribution to the psychology of language. *Psychological Record* 3:310-343.
- Brunner, E. J. (1985). *Grundfragen der Familientherapie - Systemische Theorie und Methodologie*. Berlin, Heidelberg, New York, Tokyo: Springer Verlag.
- Bucci W (1988) Converging evidence for emotional structures: Theory and method. In: Dahl H, Kächele H, Thomä H (Eds) *Psychoanalytic process research strategies*. Springer-Verlag, Berlin Heidelberg New York, S 29-49.
- Busemann A (1925) *Die Sprache der Jugend als Ausdruck der Entwicklungsrhythmik*. Jena.
- Cierpka M, Ohlmeier D, Schaumburg C (1980) Personalpronomina als Indikatoren für interpersonale Beziehungen in einer psychoanalytischen Gruppentherapie. *Psychother Med Psychol* 30:212-217.
- Cierpka M, Ohlmeier D, Schaumburg C (1983) Die Veränderungen im Gebrauch von Personalpronomina während einer psychoanalytischen Gruppentherapie. *Gruppenpsychotherapie Gruppendynamik* 18:205-216.
- Clippinger J (1977) *Meaning and discourse: A computer model of psychoanalytic speech and cognition*. Johns Hopkins Univ Press, Baltimore.
- Dahl H (1972) A quantitative study of a psychoanalysis. In: Holt RR, Peterfreund E (Hrsg) *Psychoanalysis and contemporary science*. Macmillan, New York, S 237-257.
- Dahl H (1974) The measurement of meaning in psychoanalysis by computer analysis of verbal context. *J Am Psychoanal Assoc* 22:37-57.
- Dahl H (1988) A theory of emotions as appetitive wishes and beliefs about their fulfillment. In: Safran J, Greenberg L (Hrsg) *Affective change events in psychotherapy*. Academic Press, New York, S .
- Eisenmann F (1973) *Die Satzkonjunktionen in gesprochener Sprache (Idiomata, Veröffentlichungen der Tübinger Arbeitsstelle 'Sprache in Südwestdeutschland' Band 2)*. Max Niemeyer Verlag, Tübingen.
- Gottschalk LA (1974) Quantification and psychological indicators of emotions: The content analysis of speech and other objective measures of psychological states. *Int J Psychiatry in Med* 5 (4):587-611.
- Gottschalk LA, Gleser GC (1969) The measurement of psychological states through the content analysis of verbal behaviour. *Calif Univ Press, Berkley*.
- Grünzig HJ, Mergenthaler E (1986) Computerunterstützte Ansätze. Empirische Untersuchungen am Beispiel der Angstthemen. In: Koch U, Schöfer G (Hrsg) *Sprachinhaltsanalyse in der psychosomatischen und psychiatrischen Forschung: Grundlagen- und Anwendungsstudien mit*

- den Affektskalen von Gottschalk & Gleser. Psychologie Verlags Union, Weinheim München, S 203-212.
- Horowitz, M. J., et al. (accepted for publication). Topics and signs: Defensive control of emotional expression.
- Hölzer M, Kächele H, Mergenthaler E, Luborsky L (in prep.) Vocabulary measures for the evaluation of therapy outcome: Studying the transcripts from the Penn Psychotherapy Project. .
- Hölzer M, Scheytt N, Pokorny D, Kächele H (1990) Das "Affektive Diktionär". Ein Vergleich des emotionalen Vokabulars von Student und Stürmer. PPmP-Diskjournal 1:.
- Hölzer M (1992) Das Affektive Diktionär Ulm. PPmP DiskJournal 3:
- Hölzer M, Dahl H, Kächele H (1991) A method for identifying frames. in prep.:
- Kächele H (1976) Maschinelle Inhaltsanalyse in der psychoanalytischen Prozessforschung (Reprint 1986). PSZ-Verlag, Universität Ulm.
- Kächele H (1983) Verbal activity level of therapists in initial interviews and long-term psychoanalysis. In: Minsel W, Herff W (Hrsg) Methodology in psychotherapy research. Lang, Frankfurt, S 125-129.
- Kächele H (1988) Clinical and scientific aspects of the Ulm process model of psychoanalysis. Int J Psychoanal 69:65-73
- Kächele H, Mergenthaler E (1983) Computer-aided analysis of psychotherapeutic discourse. In: Minsel R, Herff W (Hrsg) Methodology in Psychotherapy Research. Peter Lang, Frankfurt, Bd 2, S 116-161
- Kächele H, Mergenthaler E (1984) Auf dem Wege zur computerunterstützten Textanalyse in der psychotherapeutischen Prozeßforschung. In: Baumann U (Hrsg) Psychotherapie: Makro-/Mikroperspektive. Verlag für Psychologie Dr. C. J. Hogrefe, Göttingen Toronto Zürich, S 223-239.
- Koch U, Schöfer G (1986) Sprachinhaltsanalyse in der psychosomatischen und psychiatrischen Forschung. Psychologie Verlags Union, Weinheim München.
- Lolas F, Mergenthaler E, Rad Mv (1982) Content analysis of verbal behaviour in psychotherapy research: A comparison between two methods. Brit J Med Psychol 55:327-333.
- Lorenz M, Cobb S (1954) Language patterns in psychotic and psychoneurotic subjects. A. M. A. Arch. Neurol. and Psychiat. 72(6):665-673.
- Luborsky L, Spence DP (1971) Quantitative research on psychoanalytic therapy. In: Bergin AE, Garfield SL (Hrsg) Handbook of psychotherapy and behavior change : An empirical analysis. Wiley, New York, S 331-368 & 426-427.
- Mergenthaler E (1986) Die Ulmer Textbank. Entwurf und Realisierung eines Textbankverwaltungssystems als Beitrag der angewandten Informatik zur Forschung in der Psychoanalyse. Springer-Verlag, Berlin Heidelberg New York Tokyo.

- Mergenthaler E (1992). Emotion/abstractness patterns as indicators of "hot spots" in psychotherapy transcripts. In: SPR (Hrsg) 23rd Annual International Meeting, Berkeley, CA, Society for Psychotherapy Research
- Mergenthaler E, Kächele H (1985) Changes of latent meaning structures in psychoanalysis. *Sprache und Datenverarbeitung* 9 (2):21-28.
- Mergenthaler E, Pokorny D (1989) Die Wortartenverteilung - Eine linguo-statistische Textanalyse. In: Faulbaum F, Haux R, Jöckel KH (Hrsg) SOFTSTAT'89. Fortschritte der Statistik-Software 2. 5. Konferenz über die wissenschaftliche Anwendung von Statistik-Software, Heidelberg 1989. Gustav Fischer Verlag, Stuttgart New York, S 512-521.
- Mergenthaler, E., Stinson, C. H. (1992). Psychotherapy Transcription Standards. *Psychotherapy Research*, 2 (1), 58-75.
- O'Dell J, Winder P (1975) Evaluation of a content-analysis system for therapeutic interview. *J Clin Psychol* 31:737-744.
- Pepinsky HB (1979) A computer-assisted language analysis system (CALAS) and its applications. ERIC Document Reproduction Service, Arlington.
- Reynes R, Martindale C, Dahl H (1984) Lexical differences between working and resistance sessions in psychoanalysis. *J Clin Psychol* 40:733-737.
- Schaumburg C (1980) Personalpronomina im psychoanalytischen Prozess. Dissertation, Universität Ulm
- Scheibe G (1991) Evaluation des ambulanten psychiatrischen Behandlungsgesprächs. Emotionale und verbale Veränderungen von Patienten und Ärzten. Ulmer Textbank ISBN 3-926002-04-2, Ulm.
- Spence DP (1968) The processing of meaning in psychotherapy: Some links with psycholinguistics and information theory. *Behav Sci* 13:349-361.
- Teller V (1988) Artificial Intelligence as a basic science for psychoanalytic research. In: Dahl H, Kächele H, Thomä H (Hrsg) Psychoanalytic process research strategies. Springer-Verlag, Berlin Heidelberg New York, S 163-177.
- Teller V, Dahl H (1981) The framework for a model of psychoanalytic inference. *Proc of 7th IJCAI* 1:394-400.
- Thomä H, Kächele H (1988) Lehrbuch der psychoanalytischen Therapie. Bd 2: Praxis. Springer, Berlin, Heidelberg, New York, Paris, London, Tokyo
- Wirtz E, Kächele H (1983) Emotive aspects of therapeutic language: A pilot study on verb-adjective-ratio. In: Minsel W, Herff W (Hrsg) Methodology and psychotherapy research. Lang, Frankfurt, S 130-135.
- Zimmer JM, Cowles KH (1972) Content analysis using FORTRAN. *J Counsel Psychol* 19:161-166

Textsorte	Verfügbar als	Anzahl
1 Beratung	Transkript, Tonträger und Video	4
	Tonträger	1
	Video	1
2 Kurztherapie	Transkript, Tonträger	153
	Transkript, Tonträger und Video	17
	Transkript	2
	Tonträger	584
	Video	314
	k.A.	5
3 Analytische Psychotherapie	Transkript, Tonträger	27
	Transkript, Video	19
	Transkript	91
	Tonträger	1484
	k.A.	14
4 Psychoanalyse	Transkript, Tonträger	1023
	Transkript	214
	Tonträger	5662
	Video	13
	k.A.	58
5 Paartherapie	Transkript	2
	Tonträger	37
6 Familientherapie	Transkript, Tonträger	31
	Transkript	28
	Tonträger	11
7 Gruppentherapie	Transkript	26
	Tonträger	140
	Video	21
8 Supportive Psychotherapie	Transkript, Tonträger	1
9 Gruppenarbeit	Transkript	3
10 Gesprächstherapie	Video	3
11 Verhaltenstherapie	Transkript, Tonträger	6
	Tonträger	32

	Video	1
12	Erstinterview-Diagnostik	
	Transkript, Tonträger	127
	Transkript, Tonträger und Video	23
	Transkript, Video	3
	Transkript	232
	Tonträger	180
	Tonträger und Video	19
	Video	73
	k.A.	8
13	Erstinterviewbericht	
	Text, Tonträger	8
	Text	365
14	Bericht über Psychotherapiestunde	
	Text	19
	Tonträger	57

Textsorte	Verfügbar als	Anzahl
15 Bericht über Psychoanalysestunde	Text, Tonträger	7
	Text	153
	Tonträger	163
16 Vorträge allgemein	Tonträger	14
	k.A.	3
18 Balintgruppe	Transkript	53
	Tonträger	89
	k.A.	3
19 Selbsterfahrungsgruppe	Transkript	46
	Tonträger	2
20 Träume	Transkript, Tonträger	36
	Transkript	91
22 Psychodiagnostik	Transkript, Tonträger	128
	Transkript	104
	Tonträger	40
23 Katamnestisches Interview	Transkript, Tonträger	41
	Transkript	15
	Tonträger	7
	Video	7
24 TAT (Thematic Apperception Test)	Transkript	183
25 Sprachprobe	Transkript	74
26 Genitische Beratung	Transkript	37
28 HIT (Holzmann-Inkblot-Test)	Text	60
29 Erfahrungsbericht	Text	19
30 Wissenschaftliche Abhandlung	Text	40
32 Kognitive Verhaltenstherapie	Transkript, Tonträger	20
	Tonträger	19
33 Supervision	Transkript, Tonträger	16
	Tonträger	5
34 Psychiatrische Behandlung	Transkript	24
36 Familien-Interview	Transkript, Tonträger	2
	Transkript	47
37 Interaktionelle Psychotherapie	Transkript, Video	28
	Transkript	1
39 Halbstandardisiertes Interview	Transkript, Tonträger	21
30. November 2003		22

	Transkript	5
	Tonträger	44
41	Katamnestisches Interview	2
	Transkript	2
42	Circuläres Fragen	19
43	Reflecting Team therap. Intervent	20
44	Diskussion	20
99	Sonstiges	37
	Transkript	129
	Tonträger	52