

Bericht zum Workshop "Transkriptionen: Standards, Tools und gesprächsanalytische Anforderungen" am 29. Juni 2001 in Bielefeld

Kirsten Schindler und Felix Sasaki

Im Zentrum des Workshops, der von Felix Sasaki und Kirsten Schindler organisiert und im Rahmen des Studienprogramms des Graduiertenkollegs "Aufgabenorientierte Kommunikation" durchgeführt wurde, stand die Idee, ein Forum für WissenschaftlerInnen aus zwei Forschungsbereichen zu bilden: der empirisch orientierten Gesprächsforschung und der Computerlinguistik beziehungsweise Texttechnologie, die sich mit der Standardisierung und technischen Realisierung von Annotationen befasst. Bei beiden Forschungsansätzen spielt die Analyse von Gesprächsdaten eine bedeutende Rolle. Den Schnittpunkt bilden dabei die Anforderungen, welche die Gesprächsforschung an Annotationen stellt, und die Daten- und Annotationsformate, welche die Computerlinguistik in verschiedenen Tools bereitstellt (zum Beispiel das XML-Datenformat).

Das Ziel des Workshops bestand darin, gesprächslinguistische Anforderungen an Transkriptionssoftware zu formulieren und computerlinguistische Möglichkeiten zu benennen, diesen zu entsprechen. Dazu wurden VertreterInnen aus beiden Bereichen eingeladen, die Transkriptionsstandards vorstellten und in der Entwicklung begriffene oder bereits entwickelte Tools und Anwendungen miteinander diskutierten. Das interdisziplinär ausgerichtete Bielefelder Graduiertenkolleg "Aufgabenorientierte Kommunikation" bot dafür einen geeigneten Rahmen, da hier StipendiatInnen aus den entsprechenden Bereichen eng zusammenarbeiten.

Der Schwerpunkt des Workshops lag auf der Präsentation von Annotations-tools zur Bearbeitung, Analyse und Dokumentation von Gesprächsdaten. Dabei bildete ein Gesprächstranskript, das im Einführungsvortrag von Elisabeth Gülich vorgestellt wurde, die gemeinsame Datengrundlage, welche die Vortragenden jeweils zum Ausgangspunkt ihrer Präsentation nahmen.

Der Workshop wurde eingeleitet von *Elisabeth Gülich* (Universität Bielefeld), die in ihrem Vortrag *Interpretieren-Transkribieren-Analysieren: Einleitende Überlegungen zur konversationsanalytischen Arbeitsweise* drei Schwerpunkte setzte. Sie machte zunächst deutlich, welche Rolle Transkribieren in der Ethnomethodologischen Konversationsanalyse spielt. Damit setzte sie auf einer abstrakteren Ebene als der Beschreibung der eigentlichen Transkriptionskonventionen an. Sie referierte in der Folge Anforderungen an Transkriptionen und -konventionen, wie sie in der Literatur beschrieben sind, aber auch in der täglichen Forschungspraxis zum Tragen kommen. Als grundlegend stellte sie dabei die Genauigkeit des Transkripts, die Entsprechung der Transkriptionskonventionen zu multimodalen Phänomenen (auditiver oder visueller Kanal) und die Interpretierbarkeit des Transkripts heraus. Im dritten Teil ihres Vortrages beschrieb sie einen Gesprächsausschnitt und das dazu entstandene Transkript. Der Gesprächsausschnitt entstammt aus einem bereits abgeschlossenen Projekt ("Kommunikation in Kontaktsituationen zwischen deutschen und französischen SprecherInnen", gemeinsam mit Ulrich Dausendschön-Gay und Ulrich Krafft) und wurde für den Workshop mit Transkriptionskonventionen aus einem neueren Projekt (DFG-Projekt "Linguistische Differenzialtypologie epileptischer und anderer anfallsartiger Störungen - Diagnostische und therapeutische Aspekte") versehen. An dem Aus-

schnitt, der zunächst auf Kassette vorgespielt wurde, zeigte Gülich verschiedene Phänomene (zum Beispiel die Darstellung von Emotionalität) und deren textuelle Spuren im Transkript. Der Ausschnitt bildete im Workshop die Grundlage und den Anknüpfungspunkt für die weiteren Vorträge (siehe zum Gesprächsausschnitt <<http://coli.lili.uni-bielefeld.de/~felix/workshop-transkription/daten.html>>).

Ulrike Gut und Jan-Torsten Milde (Universität Bielefeld) berichteten in ihrem Vortrag *Prosodische Transkription von L2 Sprache und Aufbau eines XML-annotierten Korpus* von einem laufenden Projekt im Bereich Prosodie und einem Korpus, das in diesem Rahmen aufgebaut wird. Das Projekt "LeaP - Learning Prosody", das von Ulrike Gut betreut wird, beschäftigt sich mit der Frage, wie die Prosodie einer Fremdsprache, hier Englisch oder Deutsch, erlernt werden kann. Untersucht werden Intonation, Betonung und Rhythmus der SprecherInnen. Der quantitative Ansatz von Gut stützt sich auf eine teilweise automatische Auswertung der Daten. In dem Projekt wird außerdem ein Phonemkorpus erstellt, das "PCorpus", welches auch für andere Fragestellungen und ForscherInnen verfügbar und über das WWW erreichbar sein soll.

Die Transkription wird zunächst mit Sprachanalyseprogrammen (ESPS/waves+, Praat) durchgeführt, wobei das Transkript mit dem Audiofile unmittelbar verbunden ist. Erreicht wird dies über Zeitmarken. Wird zum Beispiel ein "e" transkribiert, so umfasst das neben dem Schriftzeichen "e" eine Zeitmarke für den Beginn und eine für das Ende im Sprachsignal. Für die L2-Analyse werden fünf Spuren ("tiers") erfasst: Phrasen, Wörter, Silben, Vokale und Intonation. In jeder dieser Spuren gibt es über die Setzung der Zeitmarken eine Verbindung zum Audiofile. Werden nicht einzelne SprecherInnen, sondern Gespräche untersucht, so erhält jeder Sprecher separate Spuren. Daneben werden "vermischte" Spuren für Geräusche, Atem, Pausen etc. eingerichtet. Dies gewährleistet einen direkten Zugriff auf einzelne Phänomene, eröffnet aber gleichzeitig auch die Möglichkeit, ihre Beziehungen untereinander zu analysieren.

Im zweiten Teil des Vortrags präsentierte Jan-Torsten Milde das Format Tax ("Time aligned signal data exchange format"), eine Architektur für ein Präsentations- und Analysetool, das die beschriebenen prosodischen Informationen abspeichert und den Zugriff darauf über einen Standard-Internetbrowser ermöglicht. Das System ist jedoch nicht nur für das "LeaP"-Projekt und das PCorpus konzipiert. Die Zielsetzung bei der Entwicklung ist, Annotationen in einem möglichst nachhaltigen Format aufzubereiten, an das verschiedene Anforderungen gestellt werden: Erstens muss es die zu untersuchenden Phänomene - bei "LeaP" prosodische Merkmale - adäquat wiedergeben können. Im Beispiel von Ulrike Gut bedeutet dies, dass die von ihr annotierten Spuren wie Silben oder Vokale auch im Tool repräsentiert werden müssen. Zweitens strebt die Kodierung der Daten ein universales Format an. Diese Anforderung trägt der Problematik Rechnung, dass viele bestehende Tools nur im Rahmen bestimmter Projekte und für bestimmte Fragestellungen entwickelt wurden und nach Ablauf dieser nicht mehr verwendet werden können. Durch ein nicht spezialisiertes Datenformat, das BenutzerInnen für ihre Zwecke anpassen können, wird die Nachhaltigkeit der Daten gesichert. Drittens müssen die Daten für Menschen und natürlich auch für Computer erfassbar sein. Ein für BenutzerInnen leicht verständliches Format ist unter Umständen für die semiautomatische Verarbeitung unbrauchbar, ein leicht zu implementierendes Format ist nicht unbedingt anwenderfreundlich. Viertens sollen die BenutzerInnen

schließlich die Möglichkeit haben, ihre bisher verwendeten Tools und daraus resultierende Daten zu erhalten. Es ist also für jedes benutzerspezifische Tool eine eigene Konvertierungsschnittstelle erforderlich.

Die von Gut und Milde angestrebte Einlösung dieser Anforderungen trennt die Repräsentation der Daten in drei Bereiche. Ausgang bilden die Daten der BenutzerInnen in einem benutzerspezifischen Format, hier ESPS/waves+ beziehungsweise Praat. Sie werden überführt in das "Taxx"-Format. Taxx dient der webgestützten Darstellung und Analyse. Es ist im XML-Format gehalten, welches durch seine zunehmende Verbreitung gute Voraussetzung für den Austausch zu anderen Daten und für die Nachhaltigkeit der Datenhaltung bietet. Aus dieser Repräsentation können schließlich andere Formate erzeugt werden, wie HTML für die Ausgabe im Browser, PDF für den Druck, eine relationale Datenbank (nötig für bestimmte quantitative Analysen) etcetera. Die im Fall der prosodischen Daten bestehende Verbindung zum Audiofile bleibt erhalten, denn die oben beschriebenen Zeitmarken sind im PCorpus integriert. Die Analysefunktion ermöglicht das Auffinden einzelner Elemente auf den verschiedenen Spuren. Zudem können komplexe Anfragen an das Korpus durchgeführt werden. Das System trennt zwischen Such- und Ausgabebereich. So können Anfragen formuliert werden wie "Suche nach Wörtern, die mit einem 'e' anfangen, und gebe die entsprechenden Sätze aus". Die folgende Grafik zeigt einen Ausschnitt aus dem PCorpus. In der linken Spalte sind die verschiedenen Ebenen aufgelistet. Im rechten oberen Fenster können Suchanfragen zu den Ebenen formuliert werden. Links darunter werden die Ergebnisse ausgegeben, rechts daneben wird das Tonsignal graphisch aufbereitet. Es kann über einen Audioplayer wiedergegeben werden.

The screenshot displays the Phonem Corpus Browser interface. On the left, a list of layers is visible, including 'eno2_CV'. The main search area shows a query: 'select a) layer CV all the events where the content is contains CVW'. Below this, search results are displayed in a table:

Layer: eno2_CV	Start	End	Length	Mid	Content
glory1	???	4.691855			CVC IMG
glory1_breaks	4.691855	4.829955			VC IMG
glory1_tones	4.829955	5.007597			V IMG
glory1_trans	5.007597	5.156328			CCV IMG
glory1_vowels	5.156328	5.267280			CV IMG
glory2	5.267280	5.392987			V IMG
glory2_CV	5.392987	5.754770			CVVC IMG
glory2_breaks	5.754770	5.975495			V IMG
glory2_syll	5.975495	6.202711			CV IMG

On the right side of the interface, there is a Winamp audio player showing a waveform visualization of the audio signal corresponding to the selected search results.

Das Taxx-System bietet bislang unter anderem eine Schnittstelle für die Programme ESPS/waves+, Praat, SyncWriter (Gesprächsannotation), Transcriber (Dialogannotation), ANVIL (Videoannotation). Das heißt, Daten aus diesen Programmen können in das Taxx-spezifische XML-Format konvertiert werden. Es

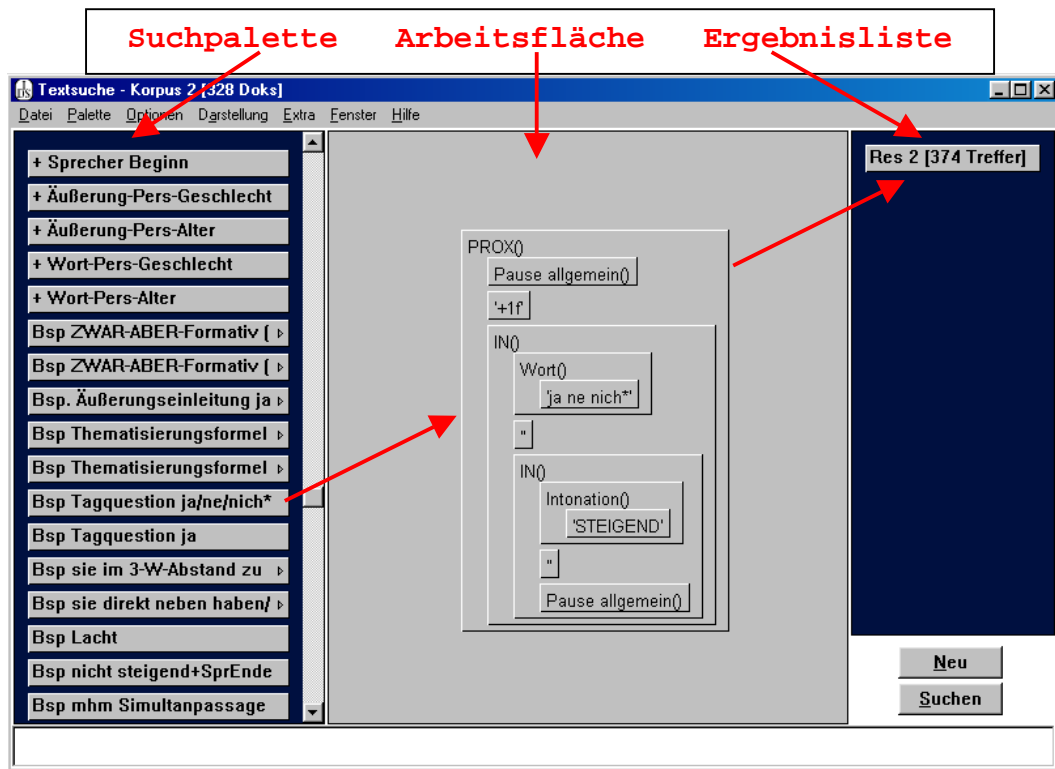
werden also keine bestimmten Annotationsformate unterstützt, sondern Schnittstellen für die Überführung in das generelle Tasx-Format.

Im Anschluss stellte *Wilfried Schütte* (Institut für Deutsche Sprache (IDS) Mannheim) *IDS-Werkzeuge für Gesprächskorpora: DIDA als Editor für Transkriptpartituren und COSMAS-II als Recherchesystem* vor. Schütte gliederte seinen Vortrag in drei Abschnitte. Im ersten Teil setzte er ähnlich wie Gülich zunächst auf einer methodischen Ebene der Beschreibung an. Schütte formulierte drei Notwendigkeiten für die Korpusarbeit: die standardisierte Erhebung von Daten, die Archivierung mit optimalem Zugriff und die definierten Zugriffsrechte. Eine weitere Notwendigkeit besteht in der Anreicherung von Korpora durch Text-Ton-Alignment (Synchronisation von Gesprächsaufnahme und Transkript) und phonetische Analyse. Eine Schwierigkeit liegt darin, dass Aussprachevarianten der Genauigkeit wegen im Transkript als solche ausgezeichnet sein müssen, die Suche nach solchen Token in einer Datenbank dadurch aber erschwert wird; zukünftig könnte dieses Problem durch eine automatische regelbasierte Lemmatisierung gelöst werden. Für die Transkription benannte Schütte eine Reihe von Anforderungen, wie zum Beispiel die Unterscheidung zwischen Sprechtext und Annotation, die Darstellung von interaktiven Bezügen im Transkript, die prosodische Notation, den Gebrauch eines Editors zur Korrektur und Annotationen für analytische Beschreibungen, die Verfügbarkeit eines integrierten Audio-Editors und die Anonymisierung von personenbezogenen Daten. Seine Forderungen bezogen sich damit sowohl auf methodische als auch auf technische Aspekte der Transkription. Schütte stellte anschließend die Gesprächskorpora am IDS und den Stand ihrer Aufbereitung vor.

Der Schwerpunkt des zweiten Abschnitts bildete die Präsentation von DIDA (DiskursDATenverarbeitung). DIDA ist ein vom IDS entwickeltes Werkzeug zur Erfassung, Pflege und Archivierung von Transkripten, das sechs Komponenten umfasst: einen Partitureditor, einen Audioeditor, eine Projektdatendank, ein Exportprogramm, ein Druckprogramm und Netzkommunikationskomponenten. Schütte stellte anschließend das von Gülich bereitgestellte Transkriptbeispiel in drei abgewandelten Formen vor: als Transkript mit Bielefelder Transkription, als Transkript mit DIDA-Transkriptionszeichen und als Transkript mit DIDA-Transkriptionszeichen nach erneutem Abhören, also retranskribiert. Schütte machte deutlich, dass mit den unterschiedlichen Konventionen auch schwerpunktmäßig bestimmte Phänomene fokussiert werden, die sich je nach Fragestellung unterscheiden können. Als wichtigste Forderung formulierte Schütte die einheitliche Nutzung der Konventionen. Gerade bei der gemeinsamen Arbeit an Korpora muss diesem Desideratum entsprochen werden, da ansonsten die Konvertierung in ein datenbanktaugliches Austauschformat (zum Beispiel SGML) stark erschwert wird.

Die Präsentation der Transkriptdatenbank COSMAS-II und entsprechender Suchoperationen stand im dritten Teil des Vortrags im Mittelpunkt. In COSMAS-II sind über 330 Diskurstranskripte eingespeist, das entspricht einer Dauer von 150 Aufnahmestunden. 20 Stunden sind davon aligniert, hier können im Transkript aufgefundene Stellen parallel abgespielt werden. Schütte führte eine Reihe verschiedener Suchoperationen (nach Worten, mit Verknüpfungen, nach Intonationsmerkmalen) durch, die er aus dem Beispieltranskript ableitete.

Die folgende Abbildung zeigt einen solchen Suchvorgang, hier nach Tag-Questions. Im Screenshot ist die Suchanfrage zweimal wiedergegeben: links in der Suchpalette in Kurzform und in der Mitte in Form einer grafischen, aus Bausteinen zusammengesetzten Abbildung. Auf der rechten Seite wird das Ergebnis der Suche angezeigt, für das gesamte Korpus erbrachte diese Suche 374 Treffer.



Die Verbindung von Transkriptionstool (DIDA) und Datenbankrecherche (COSMAS-II) hat zum Ziel, Gesprächsphänomene neu zu bestimmen, Muster in Gesprächen über eine größere Anzahl von Ausschnitten zu erkennen und neue Interpretationen verfügbar zu machen. Zwei Probleme zeigten sich im Verlauf des Vortrags und der anschließenden Diskussion. Die Transkription nach DIDA ist zeit- und arbeitsaufwendig und bedarf einer Reihe technischer Ressourcen, die in der alltäglichen Arbeitsroutine so nicht ohne Weiteres zur Verfügung stehen. Die Konventionen von DIDA werden im IDS außerdem zur verbindlichen Norm erklärt, obwohl sie von einer Reihe anderer Transkriptionskonventionen flankiert werden (wie zum Beispiel GAT). Besonders schwierig und problematisch gestaltet sich vor allem das Austauschformat. Bisher sind alle Tools und ihre Nutzung ans IDS gebunden, es besteht keine Möglichkeit des Datenaustausches mit Daten anderen Ursprungs. Weder die Datenbank noch DIDA können also bisher von einer größeren Öffentlichkeit (problemlos) genutzt werden. Dies gilt insbesondere für DIDA mit seinen proprietären Notationskonventionen und Dateiformaten. Transkripte werden nach COSMAS-II über ein TEI/SGML-Austauschformat überführt, das sich zwar fest an den DIDA-Notationskonventionen orientiert, allerdings für den Datenaustausch flexibler ist: Über TEI/SGML soll künftig ein Im-

port anderer Korpora nach COSMAS-II und ein Export in andere Datenbanken möglich sein; außerdem ist eine Internetversion des COSMAS-II-Client geplant.

Thomas Schmidt (Universität Hamburg) begann seinen Vortrag *EXMARaLDA - ein System zur Diskurstranskription auf dem Computer* mit einer Vorstellung von typischen Problemen, wie sie sich bei der Beschäftigung mit Transkriptionstools stellen und wie sie auch schon mehrfach im Rahmen des Workshops diskutiert worden waren: Die Projektbezogenheit von Formaten und Tools erschwert ihre Wieder- und Weiterverwendung. Dies wiegt umso schwerer, als sich zukünftige Anwendungsmöglichkeiten oft gar nicht erahnen lassen.

Um diesem Dilemma begegnen und detaillierte Anforderungen an ein Tool entwickeln zu können, wies Schmidt auf verschiedene Parameter hin, die für Transkriptionen von Bedeutung sind. Zunächst kann die Darstellung der Daten nach graphischen Kriterien unterteilt werden, etwa Darstellung in Partiturförm, in Spalten oder vertikal. Wünschenswert wäre daher für ein Transkriptionstool die potentielle Überführbarkeit zwischen den Repräsentationen. Dies trifft auch auf die verwendeten Datenformate zu, zum Beispiel HTML für die Darstellung im WWW, RTF für die Verarbeitung in einem Textprogramm etcetera.

Neben den Problemen der Datenhaltung beschäftigte sich Schmidt mit der Dateneingabe, -verarbeitung und -auswertung. Je nachdem, welche (graphische) Darstellungsform für die Daten bevorzugt wird, bieten sich unterschiedliche Eingabemethoden an, etwa ein Partitureditor, ein normaler Texteditor wie Word oder spezielle Programme, welche die Annotation multimedialer Daten (Video / Audio) ermöglichen. Da in dem SFB "Mehrsprachigkeit", in dessen Rahmen Schmidt tätig ist, auch außereuropäische Daten (zum Beispiel Japanisch) verarbeitet werden, müssen multilinguale Eingabeschnittstellen realisiert werden. Die Auswertung schließlich umfasst automatische Annotation, zum Beispiel unter Verwendung (morphologischer) Tagger, die Suche nach Vorkommen bestimmter Phänomene in den Daten, und schließlich eine statistische Auswertung.

Um dieser Vielzahl von Parametern zu genügen, verlässt Schmidt die Ebene der konkreten Verwendung von Transkriptionsdaten. Er konzentriert sich auf die Frage, welche Informationen beim Transkribieren kodiert werden und entwickelt daraus ein formales Modell zur Diskurstranskription. Grundlegend für dieses Modell ist die Unterscheidung und Erfassung zweier Ebenen. Jede sprachliche Einheit enthält eine Wiedergabe auf einer Zeitachse, nämlich ihren Beginn und ihr Ende. Die sprachlichen Einheiten können verschiedenste Elemente umfassen wie zum Beispiel Turns, einzelne Äußerungen, Wörter, Phoneme etcetera. Allerdings ergeben sich in Abhängigkeit von wissenschaftlichen Prämissen unterschiedliche sprachliche Einheiten, und daraus folgen unterschiedliche Transkriptionskonventionen. Unternehmen wie die TEI ("Text Encoding Initiative"), die unter anderem Konventionen für die Annotation von Gesprächen aufstellt, haben versucht, grundlegende (sprachliche) Einheiten wie Wörter verbindlich festzulegen. Es hat sich jedoch gezeigt, dass selbst diese hinterfragt werden können. Für Schmidt ergibt sich daraus, dass ein universelles Transkriptionssystem nicht festlegen darf, welche sprachlichen Einheiten zu erfassen sind, sondern nur wie: Die einzige Vorgabe, die er den BenutzerInnen macht, besteht in der beschriebenen Abbildung zeitlicher auf sprachliche Strukturen, sogenannten Annotationsgraphen (Bird/Lieberman 2001).

Das System "EXMARaLDA" (EXtensible MARkup Language for Discourse Annotation), das Thomas Schmidt am SFB "Mehrsprachigkeit" entwickelt, beruht auf dieser Abbildung (vgl. dazu auch auch den Beitrag in dieser Ausgabe). Es unterscheidet dabei drei Ebenen, die unter Verwendung des Datenformates XML erfasst werden. Auf der Ebene der Basistranskription wird lediglich die zeitliche Struktur des Diskurses beschrieben. Diese Information ist ausreichend für eine Darstellung als Partitur. Für eine vertikale Darstellung wird zusätzlich bestimmte sprachliche Information benötigt - diese kann auf der Ebene der Listen-Transkription kodiert werden, die somit auch einen Anschluss für viele vorhandene Annotationsstandards wie TEI oder GAT liefert. Die dritte Ebene ist die der segmentierten Transkription. Hier können beliebige sprachliche und zeitliche Strukturelemente nebeneinander kodiert werden. Verschiedene Darstellungsweisen können aus der beschriebenen Struktur automatisch errechnet werden. Da dieses System auf keiner spezifischen linguistischen Theorie beruht und in einer weitgehend plattformunabhängigen Sprache (JAVA) implementiert ist, können so die Daten sehr flexibel weiterverarbeitet werden.

Um die Funktionsweise von EXMARaLDA zu demonstrieren, stellte Schmidt einen typischen Ablauf vor, wie das System eingesetzt werden kann. Eine Transkription wird mit einem Partitureditor erstellt und als Basistranskription im XML-Format abgespeichert. Aus ihr können zum Beispiel Partiturdarstellungen in RTF zur Weiterbearbeitung in Word berechnet werden. Die Basis-Transkription kann nun automatisch, zum Beispiel nach Äußerungen und Wörtern, segmentiert werden. Das Ergebnis dieser Segmentierung wird in einer segmentierten Transkription - ebenfalls in XML - kodiert. Die segmentierte Transkription wird zusätzlich annotiert und in einer Datenbank gespeichert. Möglich sind nun weitere automatische Überführungen in eine Listen-Transkription oder eine vertikale Darstellung in HTML. Die folgenden Ausschnitte zeigen die Basistranskription und eine graphische Darstellung der segmentierten Transkription.

```

- <basic-transcription>
- <head>
- <meta-information>
  <project-name>EXMARaLDA Beispiele für die Homepage</project-name>
  <transcription-name>Tropfsteinhöhle</transcription-name>
  <referenced-file url="helge.wav" />
+ <ud-meta-information>
  <comment>Sketch von Helge Schneider, mit zahlreichen Beispielen für "ins Wort fallen". Die non-verbale Kommunikation wurde frei "hinzugedichtet".</comment>
  <transcription-convention>keine</transcription-convention>
</meta-information>
+ <speakertable>
</head>
- <basic-body>
  <common-timeline>
  - <tier id="TIE0" speaker="SPK0" category="v" type="t">
    <event start="T0" end="T59">Dreizehn Mann auf des toten Mannes Kiste.</event>
    <event start="T59" end="T1">((schnüffelt)) Oh, wat stinkt dat wieder in der Tropfsteinhöhle.</event>
    <event start="T1" end="T2">Oh, da kommt</event>
    <event start="T2" end="T60">einer.</event>
    <event start="T60" end="T3">Kommt noch</event>
    <event start="T3" end="T61">einer.</event>
    <event start="T61" end="T4">((hustet)) Wa/?</event>
    <event start="T5" end="T6">Oh, dat kann ja nich wahr sein.</event>
    <event start="T9" end="T8">Ja, ((Pause)) hallo</event>
    <event start="T8" end="T62">hallo Erwin.</event>
  
```

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22						
KLA [event] :	Dreizehn Mann auf des toten Mannes Kiste.														((schnüffelt)) Oh, w														
[utterance] :	Dreizehn Mann auf des toten Mannes Kiste.														((schnüffelt)) Oh, w														
[word] :	Dreizehn	Mann	auf	des	toten	Mannes	Kiste	((schnüffelt))	Oh	w																			
[turn] :	Dreizehn Mann auf des toten Mannes Kiste.														((schnüffelt)) Oh, wa														
	0	1																											
KLA [event] :	Thirteen men on the dead man's chest.														((sniffs)) Oh, what a smell here in the ca														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22						
KLA [event] :															zeigt auf sein Bein														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22						
ERW [event] :				((hustet)).			Och nee, dat jiwet ja nich.		Hallo Klaus, w																				
[utterance] :				((hustet)).			Och nee, dat jiwet ja nich.		Hallo Klaus, w																				
[word] :				((hustet))			Och nee	dat	jiwet	ja	nich	Hallo	Klaus	w															
[turn] :				((hustet)).			Och nee, dat jiwet ja nich.		Hallo Klaus, w																				
	0	1	2	3	4	5	6	7	8	9	10																		
ERW [event] :				((coughs))			Oh no, I don't believe it.		Hello Klaus, what are you do																				
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19									
ERW [event] :																					zeigt auf die Stalaktiten								
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
[event] :																						audio1							audio2

Die Loslösung von spezifischen Annotationskonventionen erklärt sich nicht zuletzt aus der Verankerung im SFB "Mehrsprachigkeit" an der Universität Hamburg. EXMARaLDA soll hier als Korpustool für Projekte zu verschiedenen Sprachen mit unterschiedlichen linguistischen Theorien dienen. Die beschriebene Architektur zielt in ihrer Generalität darauf ab, in einer solchen Umgebung unterschiedlichsten Anforderungen gerecht zu werden.

Im Verlauf des Workshops ließen sich verschiedene Tendenzen hinsichtlich des Verhältnisses von gesprächsanalytischen Fragestellungen zu Annotationssystemen auf der einen und Tools auf der anderen Seite feststellen. Deutlich wurde, dass sich die Entwicklung von projektspezifischen Annotationssystemen und eigens darauf abgestimmten technischen Realisierungen wegbewegt. Die neuen Tools erheben einen Anspruch auf möglichst generelle Einsetzbarkeit. Dabei ist ein zentrales Anliegen, die einmal erstellten Daten zu erhalten, unabhängig davon, ob sie aus anderen Projekten stammen oder mit anderer Software erstellt wurden. Die damit verfolgten Ziele liegen neben der erhöhten Datensicherheit auch in der Datenvervollständigung, denn so können auch Daten integriert werden, die bislang noch nicht elektronisch, sondern zum Beispiel in Papierform, vorliegen. Die Daten gewinnen aber auch selbst an Qualität. Erstens können Daten aus verschiedenen kommunikativen Kanälen (auditiv, visuell) gleichzeitig erfasst und in der

Analyse zueinander in Bezug gesetzt werden. Zweitens gewinnt dieser Prozess auch zunehmend an Genauigkeit und Granularität. Es werden Analysen möglich, die die Verschriftlichung von Gesprächen an Ton- und Videosignal koppeln und Beziehungen zwischen unterschiedlichen sprachlichen Ebenen (einzelne Laute bis ganze Gespräche) quantitativ (leichter) erfassbar machen.

Der Generalitätsanspruch neuerer technischer Realisierungen geht einher mit einer Trennung von projekt- oder theoriespezifischen Annotationskonventionen (CHAT, GAT, TEI) von den technischen Tools. Die BenutzerInnen sollen nicht nur in die Lage versetzt werden, Daten aus für sie nötigen Programmen zu nutzen, sondern auch ihre eigenen Konventionen im generell einsetzbaren Tool verwenden zu können. Damit entfällt zumindest technisch der Zwang, sich auf einheitliche Annotationskonventionen zu beziehen.

Zuletzt noch ein Hinweis zu Aktivitäten, die sich in der Folge des Workshops ergeben haben. Vortragende und TeilnehmerInnen haben im Anschluss an den Workshop einen Arbeitskreis gebildet, der die Entwicklungen von Tools in den verschiedenen Bereichen und Forschungsschwerpunkten an der Universität Bielefeld koordiniert. Die Mitglieder des Arbeitskreises und ihre jeweiligen Tools verfolgen dabei verschiedene Schwerpunkte, wie Dialogannotation, multimodale Annotation etcetera. Neben dem Austausch und der Koordination der einzelnen Aktivitäten, zielt der Arbeitskreises auf die Abstimmung des verwendeten Datenformats und damit die Schaffung von Austauschmöglichkeiten der Daten einzelner Projekte.

Literatur

Bird, Steven / Liberman, Mark (2001): A formal framework for linguistic annotation. In: *Speech Communication* 33 (1,2), 23-60.

Kirsten Schindler und Felix Sasaki
Universität Bielefeld
Graduiertenkolleg "Aufgabenorientierte Kommunikation"
Fakultät für Linguistik und Literaturwissenschaft
Postfach 100131
D-33501 Bielefeld
kirstenschindler@yahoo.de; felix.sasaki@uni-bielefeld.de

Veröffentlicht am 5.10.2001

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.