

Annotationsstrukturen in Transkripten. DV-technische Strukturanforderungen für Annotate exemplifiziert an EXMARaLDA

Wolfgang Schneider

Abstract

Dieser Text richtet sich an Sprachwissenschaftler, die für die Analyse gesprochener Sprache Transkriptionen benötigen und dabei computergestützte Werkzeuge einsetzen, um den hohen Aufwand des Transkribierens zu reduzieren, die Qualität der Transkriptionen zu verbessern und den Zugang zum Sprachmaterial für die Analyse zu vereinfachen.

Im ersten Teil des Artikels werden aus äußerlichen, formalen Phänomenen verbreiteter Transkriptionskonventionen einige der beim Transkribieren implizit verwendeten inneren Strukturen erschlossen. Trotz ihrer Wichtigkeit wurden Formalia der Annotationen nur selten beachtet. Eine Liste der Beurteilungskriterien für Transkriptionswerkzeuge, die möglichst allgemein einsetzbar sein sollen, wird damit um wesentliche Merkmale erweitert.

Der zweite Teil des Artikels befasst sich mit dem EXMARaLDA-System. Dieses neu entwickelte Werkzeug wird exemplarisch daraufhin untersucht, inwieweit es die so häufig verwendeten und daher unverzichtbaren strukturierenden Elemente unterstützt, die im ersten Teil der Analyse abgeleitet wurden.

Schlagworte: Formale Strukturen in Transkripten, Annotationen, Segmente, Transkriptionskonventionen, Transkriptionseditor, EXMARaLDA.

English Abstract

The article addresses linguists who need transcriptions of spoken language as a means of linguistic analysis and therefore apply computer-based tools in order to somewhat reduce the effort of transcribing, improve the quality of transcriptions, and simplify access to the linguistic material.

The first part of the article will derive from overt formal phenomena of widespread rules of transcription some of their implicit internal structures. Despite their functional importance, these formal aspects of annotations are rarely brought into focus. This study expands the list of criteria by relevant characteristics with which to evaluate transcription tools claiming universal employability.

The second section of the article gives an analysis of the EXMARaLDA-system as a case in point. This recently developed tool will be assessed according to the question in how far it supports the widespread, thus indispensable structural elements derived in the first part of the study.

Keywords: Formal structures of transcripts, annotations, segmentation, rules of transcription, transcription editor, EXMARaLDA.

1. Strukturen von Annotaten in Transkripten
 - 1.1. Transkription nach verschiedenen Konventionen
 - 1.1.1. Beispiel nach den Konventionen des Freiburger Korpus
 - 1.1.2. Beispiel nach IDS-DIDA
 - 1.1.3. Beispiel nach GAT
 - 1.1.4. Beispiel nach HIAT (realisiert mit HIAT-DOS)
 - 1.1.5. Beispiel nach CLAN
 - 1.2. Zu annotierende Einheiten
 - 1.2.1. Einheiten im Zeitablauf
 - 1.2.2. Einheiten im Transkripttext
 - 1.2.3. Gliederungen und Positionsangaben in Transkriptionen
 - 1.3. Darstellungsmittel zum Annotieren
 - 1.3.1. Gleichzeitigkeit
 - 1.3.2. Partiturzeilen
 - 1.3.3. Intonation
 - 1.4. Der Einfluss der Transkriptionswerkzeuge auf die Konventionen
 - 1.5. Verdichtung
 - 1.6. Austauschbarkeit
 - 1.6.1. Austauschbarkeit zwischen Transkribent und Lesern aufgrund sichtbarer Merkmale
 - 1.6.2. Austauschbarkeit zwischen Computern aufgrund interner Strukturen
 - 1.6.3. Austauschbarkeit "mit der Zukunft" aufgrund "ewiger" Formate?
 - 1.7. Wiederverwendbarkeit und Verallgemeinerbarkeit
 - 1.8. Globale Annotationen ("Transkriptionsköpfe")
 - 1.9. Zwischenergebnis
2. Zum Konzept von EXMARaLDA
 - 2.1. DV-technische Merkmale
 - 2.2.1. EXMARaLDAs DV-Technik
 - 2.1.2. Export von Transkripten
 - 2.1.3. Exemplarische EXMARaLDA-Daten
 - 2.2. Globale Annotationen in EXMARaLDA
 - 2.3. Strukturtypen für Annotate in EXMARaLDA
 - 2.3.1. Annotierende Textattribute (Strukturtyp A)
 - 2.3.2. Intralineare Annotationen (Stukturtyp B)
 - 2.3.3. Annotationen von Segmenten in parallelen Zeilen (Strukturtyp C)
 - 2.3.4. Annotation mit Anfangs- und Ende-Symbolen (Strukturtyp D)
 - 2.4. Intonation
 - 2.5. Volltext-Recherche, Lemmatisierung, Zerlegung und Bindungen
 - 2.6. Positionierung und Navigation in EXMARaLDA-Transkripten
 - 2.7. Integration von Medien
 - 2.8. Übernahme vorhandener Transkripte in EXMARaLDA
 - 2.9. Anpassung von EXMARaLDA an die Konventionen oder umgekehrt?
 - 2.10. Fazit zu EXMARaLDA
3. Nachbemerkung
4. Literatur
5. URLs
6. Kommentiertes Demo-Transkript im EXMARaLDA-Format

1. Strukturen von Annotaten in Transkripten

Beim Verschriften gesprochener Sprache für die sprachwissenschaftliche Analyse wird seit langem darauf geachtet, dass dies - anders als bei der bloßen Mitschrift oder beim Diktat - in systematischer Weise erfolgt. Details dieser Systematiken sind in *Transkriptionskonventionen* festgelegt. Sie werden teilweise aufgrund projektspezifischer Anforderungen entwickelt; einige haben sich mit zunehmender Erfahrung bei der Erstellung von Transkriptionen weiter verbreitet. Sie enthalten Empfehlungen oder Vorschriften für das Erscheinungsbild von Transkriptionen im Ganzen, aber auch für Details der Schreibweise der zu notierenden Phänomene.

Das Gesprochene, die Aktivitäten der Gesprächsteilnehmer und andere Ereignisse, die das Gesprochene beeinflusst haben (könnten) oder begleiten, werden in Transkripten notiert und deskriptiv, teilweise analytisch, *annotiert*. Erst durch die mehr oder weniger präzisen Vorgaben, die ein Transkriptionssystem zur inhaltliche Funktion und formalen Gestaltung macht, werden aus beiläufigen, das Transkript begleitende *Notizen* systematische *Annotationen*. Sie betreffen

- das Gesprochene in seiner verschriftlichten Form, kurz den *Text* (z.B. Kommentare, Übersetzungen, Alternativlautungen bei nicht eindeutig verstandenen Textstellen, Kennzeichnung von Reparaturen, grammatische Analysen zu einzelnen Texteinheiten),
- lautliche Phänomene, kurz den *Ton* (z.B. Prosodie, Lautstärke, Nebengeräusche),
- *nicht-verbale Aktivitäten* der Beteiligten (z.B. Gestik, Mimik, Handlungen während des Sprechens) oder auch
- die *Gesprächssituation insgesamt* (z.B. Einflüsse aus der Umgebung, die zum Verständnis des Gesprächsverlaufs relevant sind, wie z.B. das Auftreten weiterer Personen oder das Vorliegen von unhörbaren bzw. unsichtbaren Umständen).

Aus einigen der für die Erstellung deutschsprachiger Transkripte verwendeten Transkriptionssysteme werden die darin vorgesehenen Typen von Annotationen hier formal systematisiert. Mit dieser selektiven Betrachtung werden die den Transkripten inhärenten Strukturen zum Annotieren verdeutlicht. Der Überblick dient dazu, einige bisher wenig beachtete Kriterien zu gewinnen, an denen sich Transkriptionswerkzeuge messen lassen müssen, wenn sie möglichst vielseitig und allgemein einsetzbar sein wollen. Auch Recherche-Werkzeuge, die dem Auffinden von Belegstellen in Transkriptionskorpora dienen, können bei der Suche nur dann effizient auf alle in den Transkripten enthaltenen Informationen zurückgreifen, wenn sie die in den Transkripten verwendeten Annotationsstrukturen berücksichtigen (können).

Die nachfolgend betrachteten Transkriptionskonventionen sind

- das System des Korpus "Grundstrukturen der deutschen Sprache" der ehemaligen Forschungsstelle Freiburg des IDS, kurz Freiburger Korpus (FKO, Projektlaufzeit: 1960-1974).¹ Sie gehören zu den älteren Transkriptionskon-

¹ Das Material wird heute in Deutschen Spracharchiv des IDS in Mannheim archivarisches gepflegt. Aus der Dokumentation des IDS unter www.IDS-Mannheim.de/DSAv/KORPORA/

ventionen. In diesen Transkripten wurden syntaktische, ablauforientierte, lexikalische und bei einigen Transkripten auch prosodische Annotationen vorgenommen. Die verwendeten Mittel zur Annotation sind formal sehr schlicht und damit technisch einfach umsetzbar. Sie entsprechen den damaligen Möglichkeiten der Datenverarbeitung; als Datenträger kamen ursprünglich wohl Lochkarten zum Einsatz. Dieses alte Korpus wird hier deshalb betrachtet, weil darin deutlich wird, wie sehr bei reichhaltiger analytischer Annotation der Transkripttext mit den Annotationen verwoben wird.

- das System der HalbInterpretativen ArbeitsTranskriptionen² (HIAT) (Ehlich/Rehbein:1976). Transkripte nach HIAT werden in Partituren geschrieben und enthalten insbesondere Annotate zur Intonation (Ehlich 1979) und zu non-verbalen Interaktionen (Ehlich 1981a+b). HIAT wurde unabhängig von bestimmten Werkzeugen definiert. Zum effizienten Transkribieren in Partituren sollte man sich aber eines spezialisierten Transkriptionseditors bedienen.³
- die Transkriptionsrichtlinien des Instituts für Deutschen Sprache für die Erfassung von Transkriptionen⁴ für die DIskursDAtenbank (IDS-DIDA). Seit Ende der 70er Jahre wurden die Transkriptionen der am IDS erstellten Gesprächskorpora nach festen Konventionen erstellt. Die Konventionen wurden den technischen Möglichkeiten entsprechend ständig weiterentwickelt. Heute werden Gesprächstranskripte am IDS meist mit dem IDS-eigenen, hausinternen Transkriptionseditor DIDA⁵ erstellt. Die Transkripte werden - ähnlich wie bei den beiden HIAT-Transkriptionseditoren - in "endlosen" parallelen Sprecherzeilen notiert, die für die Präsentation zu Partituren umgebrochen werden.

FR_DOKU.HTM: Im Rahmen des Projektes Grundstrukturen der deutschen Sprache hatte die Forschungsstelle Freiburg des IDS die Aufgabe, grammatische und stilistische Besonderheiten der gesprochenen Standardsprache zu beschreiben. Dazu wurde ein umfangreiches Tonarchiv angelegt. Aus diesem umfangreichen Archiv wurde das Freiburger Korpus von ca. 500.000 Wörtern transkribiert. Zu diesem Korpus wurden die grammatischen und stilistischen Analysen der Besonderheiten gesprochener Sprache durchgeführt, die unter anderem auch Aussagen über einen Zusammenhang zwischen redestellativen Bedingungen und dem Einsatz spezifischer Ausdrucksmittel ermöglichen sollten. Bearbeitet wurden die Bereiche "Konjunktiv und Modus", "Passiv", "futurische und präsentische Tempora", "Vergangenheitstempora", "Morphologie" und "Wortumfang". Konventionen unter www.IDS-Mannheim.de/DSAv/KORPORA/FR/FR_KONV.HTM.

² Siehe unter www.DaF.Uni-Muenchen.de/HIAT

³ Es stehen speziell für HIAT zwei Editoren zur Verfügung: HIAT-DOS (für DOS-Windows-Rechner) und syncWRITER (für Apple-Macintosh-Rechner) Zum HIAT-DOS-Programm siehe www.DaF.Uni-Muenchen.de/HIAT. Das syncWRITER-Programm wird seit Jahren nicht mehr weiterentwickelt und wird mittlerweile vom Hersteller med-i-bit (Hamburg) auch nicht mehr vertrieben. Das im zweiten Teil dieses Textes besprochene EXMARaLDA wurde als Nachfolger für syncWRITER konzipiert.

⁴ Siehe die "Transkriptionsrichtlinien für die Eingabe in DIDA" unter www.IDS-Mannheim.de/prag/dida/dida-trl.pdf

⁵ DIDA wird derzeit nur auf Unix/Linux-Rechnern des IDS eingesetzt. Die Schreibweisen für literarische, deutschsprachige Transkripte nach den DIDA-Konventionen sind so einfach gestaltet, dass die Konventionen - abgesehen von den Pfeil-Zeichen - theoretisch auch auf jeder Schreibmaschine realisiert werden könnten. Mit DIDA lassen sich aber auch kyrillische und phonetische Zeichen verwenden. Ein kleiner Ausschnitt aus dem Unicode-Zeichenvorrat steht für DIDA zur Verfügung. Vor allem aber lassen sich die mit dem DIDA-Editor erstellten Transkripte für verschiedene Zwecke weiter nutzen, u.a. im IDS-eigenen Retrieval-System COSMAS-II.

- das System der GesprächsAnalytischen Transkriptionen⁶ (GAT, Selting et al.: 1998). Es wurde definiert, um damit mehrere Transkriptionsverfahren zu vereinheitlichen. Sie waren recht ähnlich, wohl weil sie aus der gemeinsamen Tradition der "conversation analysis" entstanden sind. Das Konzept sieht Mindeststandards für ein Basistranskript vor. Bei Bedarf kann es für weitere Analysen um differenziertere Annotationen z.B. zur Prosodie ergänzt werden und wird dann zum Feintranskript. Es werden keine Vorgaben über die Werkzeuge gemacht, mit denen die Transkripte zu erzeugen sind. Wegen der wenigen Sonderzeichen reichen fast alle Standardtextverarbeitungsprogramme dazu aus. An die GAT-Transkripte werden explizit keine DV-technischen Anforderungen gestellt.
- das System der Codes for the Human Analysis of Transcripts (CHAT), Computerized Language ANalysis (CLAN) des CHILd Language Data Exchange System (CHILDES).⁷ Es wurde konzipiert für Transkriptionen zur Erforschung des Spracherwerbs, wird jedoch auch für andere Transkriptionen verwendet.

Den Formenreichtum der in deutscher Sprache veröffentlichten Transkriptionen und insbesondere auch der dabei verwendeten Transkriptionskonventionen kann man im Repertorium 'Deutsche Transkripte 1950 bis 1995'⁸ (Ehlich/Glas 2000) erkennen. Der Überblick über 'Notation systems in spoken language corpora' (Lenk 1999) enthält zahlreiche, international ausgewählte Transkriptionsbeispiele in Englisch und in anderen Sprachen. Auch anhand dieser Konventionen könnte man die nachfolgenden Ausführungen zu Annotationsstrukturen in Transkriptionen belegen. Die prinzipiellen Schwierigkeiten von Transkriptionen, die mit ihnen verbundenen Analyseziele und einige der konkurrierenden Konzepte die Probleme theoretisch und praktisch zu handhaben, werden in 'Aufbau und Gestaltung von Transkriptionssystemen' (Redder 2001) dargestellt.

Die vorwiegend für phonetische Untersuchungen konzipierten Korpora und deren Transkriptionskonventionen⁹ sollen hier nicht näher betrachtet werden, denn sie unterscheiden sich von denen der Gesprächsforschung unter anderem dadurch, dass dort weniger Annotationen zum Transkripttext von Interesse sind, sondern vielmehr transkriptähnliche Texte als Annotationen zum visualisierten Sprachsignal (z. B. Oszillogramm oder Sonagramm) fungieren. Die andere Schwerpunktsetzung wird auch an den verwendeten Werkzeugen erkennbar.¹⁰

⁶ Siehe unter www.fbfs.Uni-Hannover.de/sdls/schlobi/schrift/GAT.

⁷ Siehe unter www.clan.org

⁸ Siehe unter www.DaF.Uni-Muenchen.de/DTR. Das Repertorium enthält eine Transkript-Bibliographie (mit 1274 Einträgen, ca. 60% von ihnen weisen das verwendete Transkriptionsverfahren explizit aus). Außerdem enthält das Repertorium eine gesonderte Bibliographie der Transkriptionsverfahren (mit 69 Einträgen).

⁹ Z.B. die Korpora des BAS (Bayerisches Archiv für Sprachsignale) und die dort entwickelte Konvention "Partitur", siehe unter www.phonetik.Uni-Muenchen.de/Bas/BasFormatsdeu.html#Partitur. "Partitur" wurde auch im Verbmobil-Projekt verwendet.

¹⁰ Z.B. wird für phonetische Untersuchungen häufig "praat" (siehe unter www.praat.org) eingesetzt. Es ist ein sehr vielseitiges Programmpaket zur Analyse von Sprachsignalen. Damit lassen sich auch Textlabel - ähnlich einem Transkript - an das Sprachsignal heften. Auch TASX (siehe Fn 48) fällt in diese Kategorie.

1.1. Transkription nach verschiedenen Konventionen

Nachfolgend wird ein (fiktives) Sprachbeispiel nach vier Konventionen transkribiert. Die Transkription soll folgende sprachliche und nicht-sprachliche Besonderheiten festhalten: Die Frage eines Sprechers mit der Hebung der Stimme am Ende der Frage, als Antwort des zweiten Sprechers eine leise gesprochene Interjektion, eine kurze Pause, gleichzeitiges Sprechen beider Sprecher ("immen" und "mor"), einen Abbruch und nachfolgende Selbstkorrektur, die Betonung auf den Silben "mor" und "fahr", die Senkung der Stimme am Ende der Aussage und eine Zeige-Geste mit dem Daumen.

1.1.1. Beispiel nach den Konventionen des Freiburger Korpus

S1: kommst du morgen mit uns mit?
 S2: / (hm) +p+
 S1: / zum Sch k+ wimmen⁹ +k .
 S2: / k+ mo:r +k gen +g+ lau (eh) fa:h⁶ ich f+ inne +f
 Stadt⁹. tttttt S1 zeigt bei "fahr ich" mit dem Daumen
 zum Fenster

1.1.2. Beispiel nach IDS-DIDA:¹¹

207 S1: kommst du morgen mit uns mi"t↑ zum
 208 S2: >hm< **
 209 S1: schw|immen↑ |
 210 S2: |mo"r |gen lau/ äh fa"h ich in=ne stadt↓
 211 K ZEIGT MIT DEM DAUMEN ZUM
 212 S2:
 213 K FENSTER

1.1.3. Beispiel nach GAT:

207 S1: kommst du morgen mit uns mIT↑
 208 S2: hm=hm, (--)
 209 S1: zum sch[wimmen]
 210 S2: [m↑or]gen LAU/ äh ↑ FAHr ich <<all> in
 zeigt mit dem Daumen zum Fenster
 211 die sta↓ dt

¹¹ Die DIDA-Ausgabe wurde nach den Richtlinien des IDS nachgebildet. Der Transkriptionseditor DIDA steht außerhalb des IDS nicht zu Verfügung.

1.1.4. Beispiel nach HIAT (realisiert mit HIAT-DOS):¹²

210	<pre> S1[Kommst du morgen mit uns mit? ((Pause 0,8s)) Zum > S2[\ / Hm, ((Pause 0,5s)) </pre>
211	<pre> S1[Schwimmen. > S2[/ Morgen lau/ äh [8 fahr ich 8] in die Stadt. [8 zeigt mit dem Daumen zum Fenster </pre>

1.1.5. Beispiel nach CLAN:¹³

```

*S1:   kommst du morgen mit uns mit?
*S2:   hm #0_5 .
*S1:   zum Sch <wimmen> [>] .
*S2:   <mo/r> [<] gen lau [////]
       äh <fa//hr ich> [%act: zeigt mit dem
       Daumen zum Fenster] in die Stadt.
    
```

Obige Beispiele verdeutlichen nur einen kleinen Teil des Formenreichtums der hierzulande etablierten Konventionen. Trotz äußerlicher Unterschiede in den Transkriptausschnitten lassen sich strukturelle Gemeinsamkeiten erkennen, von denen später noch die Rede sein soll.¹⁴

1.2. Zu annotierende Einheiten

Mit Annotaten in Transkripten werden Ereignisse teils im Zeitablauf, teils ohne expliziten Zeitbezug eng auf den Transkripttext bezogen, notiert bzw. kommentiert.

¹² Diese Partitur wurde mit HIAT-DOS erstellt, als RTF-Datei exportiert, mit dem Konvertierungsprogramm H2_UNIC in RTF mit Unicode umgewandelt und dann in Word übernommen.

¹³ CLAN-Transkripte "kehrten ihr Innerstes nach außen", d.h. alle Kodierungen sind reine Textkodierungen, die der Transkribent selber explizit vornehmen muss und die im Transkript unverändert lesbar bleiben. Solche Transkripte können deshalb auch mit jedem leistungsfähigen Texteditor erstellt werden. Der CLAN-Editor ist nur erforderlich, um solche Transkripte effizienter zu nutzen, u.a. um synchron zum Transkript auf die digitalisierte Ton- oder Videoaufnahme zuzugreifen.

¹⁴ Ähnlich argumentieren auch Bird und Liberman (2001). Als Motivation und als Beleg für ihr Modell der Annotationsgraphen bieten sie einen ebenfalls an den Annotationsstrukturen orientierten Überblick über sechs Projekte und deren Transkriptionsdaten. Von den dort besprochenen dürften für deutschsprachige Transkriptionen nur "CHAT" aus dem Childes-Projekt und "Partitur", das für die Belange von Phonetikern am BAS entwickelt wurde, eine Rolle spielen.

1.2.1. Einheiten im Zeitablauf

Um Ereignisse oder das Gesprochene selbst in der richtigen zeitlichen Abfolge festzuhalten, werden Strukturen verwendet, mit denen *Zeitpunkte* (z.B. Sprecher beginnt zum Zeitpunkt t_n einen Redebeitrag, Tür fällt zum Zeitpunkt t_n ins Schloss) oder *Zeiträume* (= von Zeitpunkt t_i bis Zeitpunkt t_j , z.B. Sprecher spricht von t_i bis t_j oder eine Sprechpause dauert von t_i bis t_j) gekennzeichnet und annotiert werden können. Ereignisse, die so schnell wieder vorbei sind, dass ihre zeitliche Ausdehnung nicht als solche wahrgenommen wird, werden eventuell nur als punktuelle Ereignisse, d.h. mit der Angabe nur eines Zeitpunktes festgehalten. Zeitpunkte und Zeiträume werden teilweise in absoluten physikalischen Größen ("Pausendauer 0,3s") oder in physikalischen Zeitkoordinaten relativ zum Beginn der Aufnahme ("0:12:34,56"), selten in absoluten physikalischen Zeitkoordinaten ("1999-12-31, 23:45:56,7") angegeben, teilweise auch nur relativ zum Höreindruck ("längere Pause") oder relativ zu anderen Ereignissen ("gleichzeitig"). Die Granularität der notierten Zeitangaben hängt entweder ab von der Genauigkeit der physikalischen Vermessung der Aufnahme (in der Regel nicht feiner als tausendstel Sekunden) oder nur vom Höreindruck des Transkribenten.

1.2.2. Einheiten im Transkripttext

Größere oder kleinere *Segmente* des Textes werden für Annotationen markiert bzw. eingegrenzt und mit Annotationen charakterisiert. Die Bezeichnung *Segment* wird hier ohne Bezug auf theoretische Konzepte zur Zerlegung gesprochener Sprache verwendet.¹⁵ Hier kommt es nur auf die formalen Aspekte der Zerlegung eines Transkripts an. Wie schon ein kurzer Blick auf einige der hier zu berücksichtigenden Transkriptionskonventionen zeigt, gibt es sehr unterschiedliche theoretische Vorstellungen darüber, wie Transkripte für die Analyse gesprochener Sprache geeignet zu segmentieren sind. Je nach Untersuchungszweck und Transkriptionskonvention werden Laute, Intonationseinheiten, Silben, Wörter, Wortfolgen, Sätze, Äußerungen bzw. komplette Redebeiträge oder noch andere Bestandteile als Segmente abgegrenzt. Teilweise lassen sich die Segmente hierarchisch einander zuordnen (z.B. Sätze bestehen aus mehreren Wörtern, eine "enthalten in"-Beziehung), teilweise nicht (z.B. sind betont gesprochene Silben Teile von Wörtern, betont gesprochene Sequenzen lassen sich aber nicht immer mittels Wörtern und deren Grenzen definieren); und verschiedene Fragestellungen führen selbst bei streng hierarchischen Segmentierungen zu so unterschiedlichen Zerlegungen eines Transkripts, dass es nur ausnahmsweise gelingen kann, die eine als eine Verfeinerung der anderen aufzufassen. Die an der Analyse der Sprache orientierten Segmentierungen passen nicht immer zu den am zeitlichen Ablauf orientierten Segmentierungen, z.B. werden für grammatische Analysen oder lexikalische Annotationen andere Einheiten abgegrenzt bzw. zusammengefasst als für phonologische oder handlungsorientierte Fragestellungen. Insbesondere bei Aufnahmen, in denen mehrere Sprecher zeitweilig auch parallel sprechen, sind zur Kennzeichnung der zeitlichen Relationen der Sprecherbeiträge weitere Abgren-

¹⁵ Eine sehr ausgearbeitete Klärung des Begriffs "Segment" findet sich bei Rehbein (1995).

zungen erforderlich, die in der Regel weder mit den grammatisch noch den phonetisch induzierten Abgrenzungen übereinstimmen.

1.2.3. Gliederungen und Positionsangaben in Transkriptionen

Transkriptionen werden in der Regel ohne gliedernde Elemente, z.B. Zwischenüberschriften, notiert. In den genannten Konventionen ist nicht vorgesehen, mit *äußerlichen Annotationen* Transkripte zu strukturieren oder Orientierungspunkte an markanten Stellen zu verankern. Annotationen dieser Art tragen wenig zur Analyse bei, erleichtern aber die Navigation in umfangreichen Transkripten. Bei längeren Aufnahmen erfolgt eine Gliederung manchmal durch Zerlegen in Teiltranskripte. Solche Einheiten werden durch die Zeiträume, die sie umfassen, eindeutig abgegrenzt. Sie lassen sich betiteln und mit weiteren globalen Annotationen (siehe dazu unter 1.8) versehen. In HIAT-Transkripten, die mit dem syncWRITER erstellt wurden, können Synchronpunkte zur Positionierung dienen. Annotationen, die sich auf absolute oder relative Zeitpunkte beziehen, bleiben auch dann verwendbar, wenn Transkriptionen verfeinert oder aus anderen Gründen später geändert werden.

Verbreitet sind verschiedene Formen automatisch erzeugter äußerlicher Annotationen für Orientierungszwecke. Beim Export werden oft die Zeilen eines Transkripts fortlaufend durchgezählt und am Zeilenanfang notiert [GAT, IDS-DIDA]. Bei HIAT-Transkriptionen werden die Partiturböcke gezählt. Diese Annotationen sind vom Umbruch abhängig, also von der verwendeten Zeilenlänge und der Schriftart. Im IDS-DIDA-Editor lassen sich beim Transkribieren Positionen mit Spaltennummern (=Anzahl der Zeichen, gezählt vom Beginn der Transkription) angeben (das ist vom Umbruch unabhängig, setzt aber fixlaufende Schrift voraus, damit die verschiedenen Sprecherzeilen synchron durchgezählt werden.).

Besonders zu beachten ist: Alle diese äußerlichen Annotationen zur Unterstützung der Navigation oder Positionierung in Transkripten ändern sich, wenn die Transkripte um Kommentare in weiteren Zeilen ergänzt werden oder wenn für weitere Annotationen beim Verfeinern der Transkription mehr Platz benötigt wird.

1.3. Darstellungsmittel zum Annotieren

In den verschiedenen Transkriptionskonventionen sind sehr unterschiedliche Techniken zum Annotieren anzutreffen. Jede dieser Techniken läßt sich formal einer der folgenden *vier Grundformen* zuordnen:

A Annotation durch *eine dem zu annotierenden Phänomen entsprechend ausgedehnte Abwandlung des üblichen Schriftbildes* des Transkripts.

Beispiele:

1 *Änderungen der Schreibweise:*

zur Kennzeichnung prosodischer Besonderheiten: Versalien für akzentuiert gesprochene Segmente "ich sprech hier **l**AUter als sonst". [z.B. IDS-DIDA, GAT].

2 *Änderungen der Schriftauszeichnung:*

zur Kennzeichnung prosodischer Besonderheiten: Unterstreichung für betont gesprochene Segmente "ich sprech hier lauter als sonst" [HIAT] oder unterbrochene Unterstreichung für gedehnt gesprochene Segmente: "eines" [HIAT]

B Annotation durch *punktueller Einfügung* von freiem oder formalisiertem Text (oder ikonischer Zeichen) in das Transkript ("intralineare Annotationen")

Beispiele:

1 Kurze *Einschübe* aus Buchstaben, Ziffern oder graphischen Schriftzeichen:¹⁶

- *Kurzer Text:*
für nicht-sprachliche Äußerungen, die in Versalien geschrieben werden: "HUSTET" [IDS-DIDA] oder in doppelte Klammern: "((hustet))" [GAT].
- *Formelhafte Kürzel:*
für eine (geschätzt) 3,5 Sekunden lange Pause: "((3,5s))" [HIAT], "*3,5*" [IDS-DIDA], "(3.5)" [GAT] oder für eine gemessene Pause: "(3.50)" [GAT].
- *Schlüssel:*
Ziffernkodes für die Intonation: "guck mal28" für Kadenz mit Halbschluss [FKO].¹⁷

2 *Diakritika über Buchstaben oder Silben:*

"h̃ " zur Kennzeichnung der Intonation dieser Interjektion [HIAT].

3 *besondere Schriftzeichen:*

"'" für auffällige Betonung: "aber ge"rn" [IDS-DIDA]
oder ":" für gedehnte Aussprache: "la::ngsa:m" [GAT, HIAT, IDS-DIDA],
oder "=" für Elisionen: "so iss=es" [IDS-DIDA],
oder "!" für extra starken Akzent: "ak!ZENT!" [GAT].

4 *graphische Symbole:*

"↑ " für das Heben der Stimme: "wohin ↑ " [IDS-DIDA, GAT]
oder für schneller gesprochene Teile ">wohin, wohin>?" [IDS-DIDA].

¹⁶ Alle Buchstaben der Schriftsprachen-Alphabete (und auch die Glyphen asiatischer Bilderschriften) und die Satzzeichen gelten hier als Schriftzeichen. Als Sonderzeichen werden mathematische Symbole, Zeichen für Parenthesen, Währungssymbole usw. bezeichnet. Und zu den graphischen Zeichen zählen Pfeil-Symbole, Semi-Blockgrafik-Zeichen für Rahmen, "smiling faces" usw. Graphische Zeichen, deren Glyphen nicht im Unicode-Arsenal enthalten sind, werden in Transkripten kaum verwendet, da sie sich nur umständlich realisieren lassen. Zur Notation der Intonation oder von Mimik und Gesten werden jedoch gelegentlich auch nicht durch Schriftnormen standardisierte ikonographische Zeichen in Transkripte eingefügt.

¹⁷ Aus dem Korpus FR: www.IDS-Mannheim.de/DSAv/KORPORA/FR/FR_KONV.HTM

C *Annotation in Zeilen parallel zum Transkripttext.* Dabei wird teilweise auch die Extension des zu annotierenden Segmentes gekennzeichnet.

Beispiele:

- 1 Kennzeichnung von Segmenten und deren Extension in einer parallelen Zeile. Der Platz, der nicht für den Text der Annotation benötigt wird, wird mit Bindestrichen aufgefüllt.¹⁸

12

RN	[ham lauter wichtige Bücher geschrieben, ich weiß.
		<i>ironisch</i> ----->
SE	[Sicher net
		<lauter

13

SE	[so wichtige wie <u>Sie</u> ein/ einmal schreiben werden, <u>weil</u>
		<i>und schneller</i> -----

14

SE	[<u>Sie</u> ham ja eine . <u>wirklich</u> große Erfahrung .. mit Macht
		----->

- 2 *Abgrenzung von Annotationen durch Klammern* (eventuell kombiniert mit anderen Hervorhebungen):

"sicher net so wichtige wie Sie ein/ einmal"

"<lauter und schneller----->"

zur Kennzeichnung der Lautstärke und des Sprechtempos [HIAT].

- 3 *Annotationen ohne exakte beidseitige Abgrenzung der Extension* (eventuell kombiniert mit anderen Hervorhebungen, z.B. Änderung der Schriftauszeichnung):

"mir san dahem jebliem und ham jegessen"

"wir sind daheim geblieben und haben gegessen"

für eine parallele Übersetzung ins Hochdeutsche [HIAT] (dabei erfolgt die Synchronisierung nicht wortweise, d.h. "jegessen" und "haben" sind übereinander positioniert, das untere Wort soll nicht als "synchrone" Übersetzung des darüberstehenden verstanden werden).

D *Annotation durch Einfügung von Symbolen (textuellen oder ikonographischen) zur Kennzeichnung des Anfangs und Endes* eines zu annotierenden Segmentes des Transkripts. Die Annotation selbst ist eventuell unmittelbar mit den Abgrenzungskennzeichen verbunden, oder sie wird in parallelen Zeilen, Spalten oder in Fußnoten notiert.

¹⁸ Transkriptausschnitt aus dem Demotranskript Nr. 1 des HIAT-DOS-Programms, Das Material wurde von Susanne Scheiter unter dem Titel "Talk im Sessel" - nächtliche Fernsehunterhaltung" für Demonstrationszwecke zur Verfügung gestellt. Siehe auch unter www.DaF.Uni-Muenchen/HIAT.

Beispiele:

- 1 *Abgrenzung* von Segmenten durch Einschleiben von Beginn- und Ende-Symbolen in den Transkripttext (hier zur Kennzeichnung von Simultanpassagen):
 S1: ja | aber |
 S2: | nein nie | mals [IDS-DIDA]
- 2 *Abgrenzung* von Segmenten durch Einschleiben von Beginn- und Ende-Symbolen, die zugleich als Träger der Annotation fungieren, in den Transkripttext (hier zur Abgrenzung von Fremd- bzw. Dialektwörtern ("f+ biste +f") oder eines erweiterten Infinitivs mit einer Sequenz simultanen Sprechens und einem Eigennamen):
 S2: k+ (ja6) +k
 S1: i+ k+ den z+ Goethe +z +k nicht zu zitieren6 +i
 [FKO]¹⁹
- 3 *Abgrenzung* von Segmenten durch Beginn- und Ende-Symbole, die orthogonal zur Leserichtung als "vertikale" Klammern fungieren (siehe oben unter 1.1.3 [GAT]).
- 4 Insbesondere Kommentare zu nicht-sprachlichen "Vorkommnissen" werden oft (nach Art eines Drehbuchs) in einer *zweiten Spalte* neben den Transkripttext geschrieben. Die Stelle oder der Bereich, worauf sich der Kommentar bezieht, wird durch Beginn- und Ende-Marken oder ähnlich wie bei Fußnoten kenntlich gemacht: "[3 ... 3]" [HIAT].

Welche der vier grundlegenden Annotationsformen verwendet wird, hängt weitgehend von der Transkriptionskonvention ab. Aber auch innerhalb eines Transkriptionssystems kann noch die Länge der Einheit, auf die sich die Annotation bezieht, und der Umfang der Annotation für die zu verwendende Annotationsform eine Rolle spielen. Z.B. enthalten FKO-Transkripte keine Annotationen der Form A oder C. In Transkripten nach HIAT oder GAT hingegen können Annotationen aus allen vier Grundformen vorkommen.

Die bewußte Abweichung von der Orthographie (s. obiges Beispiel "lAUter") wird oft nur als besondere Schreibweise der Transkription aufgefaßt. Versalien mitten im Wort eines Transkripttextes erfüllen aber eine abgrenzende und kennzeichnende Funktion. Mit dieser Schreibweise wird also eine Annotation vorgenommen. Die Extension der Annotation kennzeichnet zugleich die Extension eines Segmentes. Eine solche *inhärente Annotation eines Segmentes* mit der Bedeutung "Betonung" wird hier selbst zum Teil der Verschriftung des Gesprochenen.

1.3.1. Gleichzeitigkeit

Zur *Kennzeichnung von Gleichzeitigkeit* werden gelegentlich besondere Techniken beim Layout von Transkripten verwendet. In den einfachsten Fällen werden an markanten gleichzeitigen Stellen Synchronisationsmarken in den Fließtext eingefügt (z.B. "|" für "hier beginnt bzw. endet das gleichzeitig Gesprochene"

¹⁹ Auszug aus der Interaktion FR030 des Korpus FR des DSAV am IDS, siehe unter www.IDS-Mannheim.de/DSAv/KORPORA/FR/FR0/FR030/FR030TRA.HTM.

[IDS-DIDA] oder "<" am Anfang und ">" am Ende jedes überlappenden Segmentes mit zusätzlichen Einschüben, die die Zugehörigkeit der Überlappungen kennzeichnen [CHAT]), oder es werden an solchen Stellen Zeilenwechsel vorgenommen und die gleichzeitig gesprochenen Sprechanteile mehrerer Sprecher in parallelen Zeilen notiert [GAT]. Da die Transkription bei häufigen Sprecherwechseln und parallelem Sprechen mehrerer Sprecher damit sehr unübersichtlich werden kann, ist bei einigen Transkriptionskonventionen generell eine Schreibung von Transkripten in horizontalen Partituren [HIAT, IDS-DIDA] vorgesehen.

Die *Partiturschreibweise* ermöglicht eine sehr präzise Notation der Gleichzeitigkeit bei verhältnismäßig geringem Platzbedarf.²⁰ In Partituren wird üblicherweise für jeden Sprecher eine *Sprecherzeile* (auch als *Spur* oder *tier* bezeichnet) für die Transkription des Gesprochenen verwendet. Dazu werden bei Bedarf weitere parallele Zeilen, die dem Sprecher zugeordnet sind, zur Kennzeichnung nonverbaler Aktivitäten oder der Intonation eingesetzt; sie bilden alle zusammen einen *Sprecherblock*. Die Blöcke aller beteiligten Sprecher bilden eine Partitur. Hinzu kommen eventuell noch weitere Zeilen - formal ähnlich wie Sprecherzeilen - zur Kennzeichnung von Ereignissen, die sich nicht auf einen einzelnen Sprecher beziehen. Für alle parallelen Zeilen gilt mehr oder weniger präzise, dass das über- und untereinander Notierte gleichzeitig auftrat.

1.3.2. Partiturzeilen

Die *vertikale Strukturierung der Partitur* in mehrere Zeilen für die eigentliche Transkription des Gesprochenen und für andere Zwecke bildet ein Gerüst zur Annotation der Transkription [HIAT, IDS-DIDA]. Die Zeilen lassen sich durch zwei Merkmale charakterisieren: *Adressierung* und *Zweck*. Jede Zeile ist entweder genau einem Sprecher zugeordnet²¹ oder der Gesprächssituation insgesamt. Und jede Zeile ist üblicherweise für Eintragungen zu genau einem Zweck vorgesehen, z.B. für die Verschriftung des Gesprochenen, für die Intonation, für nonverbale Kommunikation, für Verknüpfungen mit der digitalisierten Tonaufnahme oder für Kommentare zum Gesprächsverlauf. Durch den Eintrag in die jeweils "zuständige" Zeile lassen sich Annotationen ohne zusätzliche Kennzeichnungen adressieren, d.h. dem richtigen Sprecher zuordnen, und gleichzeitig für den vorgesehenen Zweck ausweisen (z.B. "nonverbale Kommunikation"). In einer HIAT-Partitur gehört z.B. jede NVK-Zeile²² immer jeweils zu genau einem Sprecher; sie steht unterhalb der Sprecherzeile für das Gesprochene.²³ Im IDS-DIDA-Editor steht unterhalb jeder Sprecherzeile eine Kommentarzeile (Kennzeichnung "K") zur Ver-

²⁰ Nur an den Stellen, an denen Äußerungen eines Sprechers zu transkribieren sind, wird auch tatsächlich Platz für diesen Sprecher benötigt; mindestens jeweils eine Zeile, mehr aber auch nicht.

²¹ Dem widerspricht auch nicht die gängige Praxis, bei Transkripten mit sehr vielen, teilweise anonymen oder unidentifizierten Sprechern (z.B. Schulunterricht oder Diskussionen) gelegentlich aus technischen Gründen eine Sprecherzeile für mehrere Sprecher zu verwenden, wenn sich die Eintragungen nicht wechselseitig stören.

²² NVK-Zeilen dienen in HIAT zur Notation der nonverbalen Kommunikation; siehe Ehlich (1981a und 1981b).

²³ In HIAT-DOS z.B. wird die Strukturierung der Partituren durch Klammerungen von Sprecherblöcken auch in der Präsentation des Transkripts hervorgehoben; dort gibt es jedoch keine Zeilen, die explizit zur Annotation des Gesprächsverlaufs insgesamt vorgesehen wären.

fügung und unter allen anderen Zeilen zusätzlich noch eine Kommentarzeile für die Gesprächssituation insgesamt (Kennzeichnung "κ&").

1.3.3. Intonation

Anhand der *Notation der Intonation* läßt sich studieren, wie unterschiedlich die Struktur von Transkripten konzipiert sein kann, um solche Phänomene zu notieren. Schon die Vorgaben der Konventionen sind strukturell nicht immer eindeutig. Intonation wird z.B. nach HIAT mit Sonderzeichen, die das Auf bzw. Ab (schräge Striche) der Stimme symbolisieren sollen, oberhalb der Verschriftung des Gesprochenen notiert.²⁴ Diese Vorgabe zielt nur auf die ikonische Präsentation der Notation. Und obwohl die Zeichen²⁵ und deren Bedeutung im HIAT-System geregelt sind, wird die Notation der Intonation bei den beiden HIAT-Editoren mit völlig disparaten technischen Strukturen realisiert.²⁶ Für das Transkribieren mit syncWRITER wird empfohlen, die Intonationszeichen mit Diakritika direkt in die Sprecherzeile zu schreiben. Für Vokale stehen Diakritika in syncWRITER problemlos zur Verfügung. Für die Kennzeichnung der Intonation von Interjektionen (z.B. "h̄m̄ ") wird dort ein spezieller Schriftfont (HIAT-Times) bereitgestellt, der die gebräuchlichen Kombinationen von Konsonanten mit Diakritika enthält. Im Transkriptionseditor HIAT-DOS stehen keine Diakritika zur Verfügung und zusätzliche Schriftarten lassen sich auch nicht hinzufügen. Aus diesem rein technischen Grund - vielleicht auch, weil der Kennzeichnung der Intonation dort konzeptionell ein anderer Stellenwert beigemessen wurde als im syncWRITER - steht bei Bedarf über jeder Sprecherzeile eine Intonationszeile zur Verfügung, in die man die Intonationssymbole über dem Transkripttext eintragen kann. Beim Erstellen des Transkripts ist die Intonationszeile sehr eng an die Sprecherzeile gebunden, damit sich die Positionierung der Sonderzeichen bei Einfügungen und Löschungen in der Sprecherzeile nicht unabsichtlich verschiebt. Intonationszeilen werden vom HIAT-DOS-Editor also nicht wie "normale" Zeilen einer Partitur behandelt. Für die besonders detaillierte Notation der Intonation sehen die HIAT-Konventionen vor, dass Tonhöhenverläufe mit kugelförmigen Symbolen - ähnlich der Notenschreibung - auf mehreren Höhenlinien (meist fünf) notiert werden. Dazu wurden zu Schreibmaschinenzeiten Halbschritte der Walze unterschieden; in den

²⁴ Welche sprachlichen Einheiten als Träger des Tons fungieren und wo genau demzufolge nach den HIAT-Konventionen die Sonderzeichen positioniert werden sollen, über Silben, über Phonemen oder nur über einzelnen Vokalen oder stimmhaften Konsonanten scheint nicht eindeutig geklärt. Die Sonderzeichen werden über einzelne Buchstaben gestellt - oft über den mittleren einer Silbe, sofern die Silbe aus einer ungeraden Anzahl Buchstaben besteht. Wie bei Diphthongen zu verfahren ist, bleibt ebenfalls etwas unklar.

²⁵ "Zeichen" im Sinne von Symbolen zur Präsentation, aber gerade nicht als Glyphen mit Kodierung im DV-technischen Sinne.

²⁶ Die in HIAT auch vorgesehene graphisch orientierte Kennzeichnung von Tonhöhenverläufen mit langen schrägen Strichen, die zugleich der Extension des zu annotierenden Textes angepaßt wurden, ist aufwändig, weil sie sich mit den existierenden Transkriptionseditoren nicht realisieren läßt. Solche *graphischen Annotationen* wurden deshalb nur selten manuell in Transkripte eingetragen und sollen hier nicht weiter betrachtet werden.

Editoren wird ersatzweise für jede Höhe eine eigene Zeile verwendet.²⁷ Bei den anderen hier betrachteten Transkriptionskonventionen ist die Intonation, so wie alle anderen prosodischen Phänomene, intralinear zu notieren - auch beim Partitur-System IDS-DIDA. Dazu sind jeweils bestimmte Sonderzeichen vorgesehen, die teilweise als freistehende Diakritika realisiert werden (z.B. Akzentzeichen "´" im GAT), jedoch keine Diakritika über Grundbuchstaben. Bei der Auswahl der Symbole zur Notation der Intonation und bei deren Positionierung (vor, hinter oder mitten im Wort oder einer Silbe) unterscheiden sich die Konventionen.

1.4. Der Einfluss der Transkriptionswerkzeuge auf die Konventionen

Bei der Konzeption von Konventionen wird immer ein Kompromiss zwischen verschiedenen Anforderungen gesucht. Eindeutigkeit und Verständlichkeit der verwendeten Symbolik, Lesbarkeit des Textes, platzsparende Notation, Reichhaltigkeit und Flexibilität zur Notation der für die Analysen relevanten Merkmale und die einfache Realisierbarkeit mit den verfügbaren Werkzeugen sind dabei zu berücksichtigen und widersprechen sich gelegentlich.

Die heute für Transkriptionen deutscher Sprache verbreiteten Systeme wurden ursprünglich für Schreibmaschinentechniken entwickelt - entweder, weil zum Zeitpunkt der Definition der Konvention noch gar keine Computer zum Transkribieren verfügbar waren oder weil die Konstrukteure sich selber und die Anwender der Konventionen nicht mit technischen Details belasten wollten, die wesentlich über die Möglichkeiten einer Schreibmaschine hinaus gingen. Die Details der Konventionen lassen sich wohl nur verstehen mit Blick auf derartige technische Einschränkungen - sowohl hinsichtlich des Layouts wie des üblicherweise verfügbaren Zeicheninventars. Dass heute mehr oder weniger geeignete Textverarbeitungssysteme (spezialisierte Transkriptionseditoren oder Standard-Textverarbeitungssysteme) zur Erstellung von Transkriptionen verwendet werden, spielte bei der Definition der Konventionen (außer bei CHAT) kaum eine Rolle.

1.5. Verdichtung

Konventionalisierte Schreibweisen in Transkripten (Formulierungen und Zeicheninventar) und konventionalisierte Gestaltungsweisen der Transkripte (Typographie und 'Layout') ermöglichen es mit ihren halbwegs präzisen Regelungen der Verwendung von Kürzeln, von Sonderzeichen und typographischen Mitteln (Unterstreichungen, Versalien, Schriftartwechsel usw.) Transkripte *platzsparend* zu annotieren. So wird der *Informationsgehalt* der Transkripte effizient *verdichtet*. Mit Kürzeln lassen sich Annotationen schneller notieren als in ausführlicher Schreibweise und trotz der höheren Komplexität lassen sich solche Annotate auch

²⁷ Da in HIAT-DOS die Anzahl der Zeilen, die für einen Sprecher vorgesehen sind, nicht frei definiert werden kann, wird dort jeweils ein ganzer, darüberliegender Sprecherblock "missbraucht". Es gibt ein technisches Konzept besondere Intonationssymbole zu definieren, die in einem Glyph die Tonhöheninformation enthält. Solche Symbole könnten - ähnlich wie in HIAT-DOS - in eine Intonationszeile über dem Text notiert werden.

schneller lesen. Es setzt jedoch eine gewisse Vertrautheit mit dem Inventar verwendeter Kürzel und Gestaltungsmittel zum Annotieren voraus.²⁸

1.6. Austauschbarkeit

Die Berücksichtigung von eigenen, projektspezifischen oder von anderen übernommenen Transkriptionskonventionen führt zu einer *Vereinheitlichung* der Transkriptionen zumindest innerhalb eines Korpus oder auch über dessen Grenzen hinweg. Die Ablösung individueller Schreibweisen und Stile einzelner Transkribenten zugunsten einheitlich geregelter Konventionen führt auch zur *Austauschbarkeit* von Transkriptionen im weitesten Sinne; sowohl zwischen Transkribenten und (anonymen) Lesern als auch zwischen DV-Werkzeugen für unterschiedliche Aufgaben. Universelle Austauschbarkeit von Transkriptionen beruht auf Regelungen, die drei verschiedene Ebenen betreffen: die Ebene der sichtbaren Oberfläche der Transkriptionen, über die sich Transkribenten und Leser verständigen und austauschen können; die Ebene der internen Datenformate der Transkriptionen, die den unmittelbaren Austausch zwischen Computern und Programmen sicherstellen; schließlich die Ebene abstrakter Datenformate, die den Austausch von Transkriptionen mit völlig fremden oder gar zukünftigen Rechnersystemen und Programmen ermöglichen (eventuell sind dazu maschinelle Konvertierungen erforderlich).

1.6.1. Austauschbarkeit zwischen Transkribent und Lesern aufgrund sichtbarer Merkmale

Transkripte sind die gegenständliche Basis dafür, sich mit anderen Interessenten (Fachleuten oder Laien) über konkrete Diskurse und Diskursabläufe zu verständigen. Dafür muss die sichtbare Oberfläche der Transkripte verständlich bzw. interpretierbar sein. Je mehr die verwendeten Konventionen Lesern und anderen Transkribenten vertraut sind, umso eher sind Transkripte austauschbar und können von anderen weiter verwendet werden. *Austauschbarkeit* bedeutet also die ganz äußerliche vom Layout und den Schreibweisen abhängige "Verständlichkeit" und "Lesbarkeit" der Transkripte für einen mit den Konventionen vertrauten Kreis.

Die oben berücksichtigten Konventionen befassen sich (abgesehen von CLAN) nur mit der sichtbaren Oberfläche von Transkriptionen. Dieses Defizit betrifft nicht nur die älteren Konventionen, die für längst abgeschlossene Korpora verwendet wurden, sondern auch die aktuellen.

Der Austausch von Transkripten, die nach diesen Konventionen erstellt wurden, kann nur aufgrund der sichtbaren Merkmale erfolgen. Der Austausch war ursprünglich auch nur in gedruckter oder fotokopierter Papierform vorgesehen. Ob auch andere Formen des Austausches solcher Transkripte praktikabel sind, ist allein durch die verwendeten Konventionen nicht entschieden.

²⁸ Welch hohe Anforderungen "Professionelles Transkribieren" insgesamt an Transkribenten und Rezipienten stellt, um "authentische" Transkripte zu erzeugen bzw. zu verstehen - mit der Kenntnis der Notationskonventionen ist es natürlich nicht getan -, wird von Redder (2002) dargelegt.

1.6.2. Austauschbarkeit zwischen Computern aufgrund interner Strukturen

Austauschbarkeit bedeutet in technischer Hinsicht die Möglichkeit, Transkripte als Dateien *weitergeben und wiederverwenden* zu können. Das gelingt nur dann fehlerfrei, wenn Hersteller und Empfänger von Transkripten identische oder aufeinander abgestimmte DV-Werkzeuge verwenden und das auch noch nach identischen technischen Regeln. Diese werden in den Transkriptionskonventionen aber normalerweise gar nicht thematisiert. Hersteller und Empfänger müssen nicht verschiedene Personen sein, auch die Möglichkeit zur Weitergabe einer Transkriptdatei an ein anderes Programm - z.B. an ein Werkzeug zur morphologischen Analyse oder ein Volltext-Retrieval-System - erfordert ein gewisses Maß an *Austauschbarkeit*. Und schließlich ist unter *Austauschbarkeit* auch die Möglichkeit der Weitergabe von Transkriptionen an andere Projekte und an die wissenschaftliche "Nachwelt" zu verstehen, die Konservierung der Transkriptionen für andere Projekte, gar spätere Generationen.

Um Transkripte so zu erstellen, dass sie austauschbar werden bzw. bleiben, müssen auch DV-technische Konventionen verabredet und eingehalten werden. DV-technische Vorgaben wurden jedoch in den oben genannten Transkriptionskonventionen nicht explizit getroffen, allenfalls peripher oder implizit.

- Die Transkripte des Freiburger Korpus liegen heute in einheitlicher Form vor, vermutlich war auch ihre Erstellung schon sehr streng geregelt. Da sie aus heutiger Sicht zudem DV-technisch recht anspruchslos sind, lassen sie sich sehr zuverlässig austauschen.²⁹
- Bei der Definition des Systems GAT, die erst zu einer Zeit erfolgte, als die Verwendung von PCs schon allgemeiner Usus war, hat man nur bei der Auswahl der zu verwendenden Glyphen für Sonderzeichen (nicht bei deren Kodierung) ein wenig auf deren allgemeine Verfügbarkeit geachtet. Wie die Transkripte technisch zu realisieren sind, bleibt erklärtermaßen außerhalb der Definitionsabsicht von GAT; die Austauschbarkeit von Transkripten nach GAT bleibt DV-technisch gesehen also dem Zufall überlassen.³⁰
- Die Transkriptionsrichtlinien des IDS wurden seit den achtziger Jahren bis heute kontinuierlich weiterentwickelt, und dabei wurde immer auch Rücksicht genommen auf die verfügbaren Werkzeuge, mit denen diese Transkripte erstellt werden konnten. Da diese Werkzeuge oft nur IDS-intern verfügbar waren, erfolgte der Austausch computerlesbarer Transkripte mit externen Interessenten allenfalls auf der Ebene der Präsentationsformate (z.B. im Rich Text Format). Transkriptionen nach der "Richtlinie für das Transkribieren mit DIDA" lassen sich auch ohne den IDS-internen spezialisierten Transkriptionseditor mit jedem leistungsfähigen Texteditor (UNICODE-Fähigkeit ist erforderlich) schreiben. Ein Austausch solcher Transkripte ist aber nicht zuverlässig.

²⁹ Es waren reine Textdateien, die bei diversen Übernahmen auf modernere DV-Systeme mehrfach umkodiert wurden. Das Deutsche Spracharchiv im IDS hat sie in einem DSAV-eigenen XML-Format archiviert und für den Zugriff aufbereitet, siehe unter www.IDS-Mannheim.de/DSAv/KORPORA/FR/FR_DOKU.HTM.

³⁰ Aus der Einleitung zu GAT: "Die Konventionen richten sich an Transkribenten, nicht an Computer; sie umfassen bisher noch keine Kriterien für die Formatierung von Dateien, die GAT folgend angefertigt werden" (Selting 1998). Bisher sind keine Bemühungen der GAT-Entwickler bekannt geworden, daran etwas zu ändern.

sig möglich, weil es bei manueller Erstellung nahezu unmöglich sein dürfte, alle Informationen, die ein DIDA-Transkript enthält, fehlerfrei einzutragen und weil es auch für den DIDA-Editor keinen Import-Konverter gibt, mit dem man Transkripte aus dem Präsentationsformat zurück in das DIDA-interne Format wandeln könnte.

- Anders ist das bei den Konventionen, die eng mit einem bestimmten DV-Werkzeug verbunden sind. Z.B. lassen sich Transkripte gemäß den HIAT-Konventionen, die schon zu Schreibmaschinenzeiten entwickelt wurden (die sich also von jedem Texteditor realisieren lassen), seit vielen Jahren auch mit Transkriptionseditoren (HIAT-DOS und syncWRITER) erstellen.³¹ Für beide gibt es detaillierte Empfehlungen, wie die HIAT-Konventionen mit dem jeweiligen Editor am besten umgesetzt werden. Der Austausch von Transkriptionen innerhalb des Anwenderkreises eines der beiden Programme ist problemlos, der Austausch zwischen beiden ist dagegen praktisch nicht durchführbar (zumindest nicht mit den Mitteln der verwendeten Programme), obwohl in allen Fällen HIAT-Konventionen verwendet wurden.
- Unter den Transkriptionskonventionen für die Gesprächsforschung ist CHAT (Codes for Human Analysis of Transcripts)³² eine Ausnahme, weil es von vornherein für die Erstellung und Auswertung am Computer konzipiert wurde. Die Transkripte sind reine Textdateien, deren formaler Aufbau präzise geregelt ist. Um Transkripte im CHAT-Format zu erstellen und zu analysieren, gibt es das Transkriptionswerkzeug CLAN.³³ Trotz der Computerunterstützung verlangen diese Werkzeuge vom Transkribenten, selber wie ein Computer zu agieren. Es werden nur wenige Hilfen angeboten, die sicherstellen, dass die Kodierung effizient vorgenommen werden kann und vor allem auch syntaktisch fehlerfrei ist. Das muss man erst sorgfältig austesten (bzw. von einem CLAN-Parser testen lassen). Der Grund dafür liegt in dem - je nach Sichtweise - "primitiven" oder auch "robusten" Format, das außer den sichtbaren Kodierungen kaum interne Kodierungen "im Hintergrund"³⁴ kennt. Als Transkribent ist man dann dafür verantwortlich, die Kodierungen der Annotationen so vorzunehmen, dass eine computergerechte Auswertung damit vorgenommen werden kann. Der Austausch gelingt nur, wenn die Annotationen syntaktisch fehlerfrei kodiert wurden, dann jedoch recht zuverlässig.

³¹ Siehe Fn 3. Mit Einschränkungen lassen sich HIAT-Transkripte auch mit EXMARaLDA erstellen; dazu siehe Kapitel 2.

³² CHILDES (Child Language Data Exchange System) befaßt sich mit dem Erstspracherwerb von Kindern in verschiedenen Sprachen und pflegt eine Transkriptionsdatenbank, die unter anderem auch deutschsprachige Korpora enthält. Das Konzept wurde von vornherein für die Austauschbarkeit der Materialien per Computer konzipiert. Siehe unter: childes.psy.cmu.edu. Seit kurzer Zeit lassen sich CHILDES-Transkripte nicht nur im CHAT-Format, sondern auch mit XML-Mitteln kodieren.

³³ Es handelt sich um einen einfachen Texteditor, der angereichert wurde mit diversen Hilfsmitteln zur Auswertung von Transkripten (Wortstatistiken etc.).

³⁴ In Transkripten, die mit digitalisierten Tonaufnahmen oder Video verknüpft sind, lassen sich allerdings Zeitmarken ein- und ausblenden.

1.6.3. Austauschbarkeit "mit der Zukunft" aufgrund "ewiger" Formate?

Austauschbarkeit kann und sollte auch *Wiederverwendbarkeit* in der Zukunft beinhalten. Die Austauschbarkeit von annotierten Transkriptionen mit anderen Wissenschaftlern in anderen Projekten, die eventuell erst sehr viel später durchgeführt werden und sich am selben Material auf ganz andere Fragen konzentrieren, kann man heute schon "sicherstellen", zumindest günstig beeinflussen. Auch wenn die künftig verfügbaren Werkzeuge und Arbeitsumgebungen heute natürlich noch völlig unbekannt sind, kann man nämlich eine spätere Übernahme in andere Systeme dadurch wesentlich vereinfachen, dass man heute bestimmte systematische Vorbereitungen trifft. Man sollte

- a.) alle Annotationen in den Transkripten explizit als solche kodieren (d.h. keine indirekten Annotate wie z.B. mit Versalien in "lAUter"),
- b.) alle Kodierungen in den Transkripten offenlegen und als reinen Text notieren (d.h. keine binären Daten oder kryptische Kodierungen verwenden, deren Bedeutung nicht unmittelbar im Klartext lesbar ist),
- c.) alle Strukturen in den Transkripten offenlegen und als reinen Text notieren (d.h. keine "heimlichen" Regeln unterstellen wie z.B. "Zeilen ohne Sprechersigle enthalten Kommentare zum Gesprächsverlauf"),
- d.) alle Kodierungen für Strukturen und Annotate dokumentieren (Lesehilfen).

Dies führt zu einer strikten Trennung von äußerlicher Präsentation (Layout) und interner Notation der Strukturen. Layout und Transkriptstrukturen sollten zudem als reiner Text notiert werden. Wenn man dann noch eine Strukturbeschreibung verwendet, die "selbsterklärend" und deren Syntax validierbar ist, kann man später mit einfachen Mitteln alle Informationen zu den Transkripten auslesen und sie in die gewünschte Form bringen. Eine Strukturbeschreibung für Transkriptionen mit den Mitteln der eXtensible Markup Language (XML) ist heute wohl der beste Weg, obige Anforderungen "im Prinzip" zu erfüllen. Wie die *Strukturbeschreibung im Detail* aussehen muss, um obige Anforderungen zu erfüllen, läßt sich *nicht eindeutig* entscheiden. Inzwischen gibt es zahlreiche Entwürfe für Transkriptformate, die mit XML-Mitteln definiert wurden und obige Anforderungen erfüllen. Eine solche Transkription ist im Anhang als kommentiertes XML-Beispiel zu finden.

1.7. Wiederverwendbarkeit und Verallgemeinerbarkeit

Das Bedürfnis oder die Notwendigkeit, einmal erstellte Transkripte auch später noch für andere, jetzt noch nicht bekannte Zwecke zu nutzen bzw. anderen zur Verfügung zu stellen, wird von den Transkriptionskonventionen in der Empfehlung berücksichtigt, dass die Transkriptionen "auf einem 'mittleren' Transkriptionsniveau gehalten werden und dann als Basis für mögliche Fein- oder Ergänzungstranskriptionen dienen können" (aus der Vorbemerkung zu den DIDA-Transkriptionsrichtlinien). Ein solches *Basistranskript* [GAT] wird auch als *Rohtran-*

skript [HIAT] bezeichnet, das vom *Feintranskript* [alle] unterschieden wird.³⁵ Für die Weitergabe stehen damit unterschiedlich umfangreich annotierte Versionen zur Verfügung. Dabei kann ein Problem entstehen: Im Zuge der Detailarbeit am Feintranskript werden gelegentlich Phänomene dechiffriert, die zu Veränderungen im Basistranskript führen müssten. Solche nachträglichen Korrekturen erfordern jedoch noch einmal den gleichen Aufwand in der zugrunde liegenden Basis wie schon im Feintranskript. Das zur Effektivierung vorgesehene Verfahren, zwei Bearbeitungsmodi von Transkripten auf dem Wege herzustellen, dass das Ausgangstranskript dupliziert und das Duplikat verfeinert wird, führt nicht zu einer Ökonomisierung - im Gegenteil. Da die zwei auseinander hervorgegangenen Transkripte technisch unabhängig voneinander existieren, erfordert das Verfahren für die konsistente Pflege und Weiterentwicklung (z. B. Korrekturen, Verknüpfung mit Ton- oder Videomaterial) beider Transkripte den doppelten Arbeitsaufwand.

Annotationen lassen sich nach dem *Grad der Verallgemeinerbarkeit* systematisieren. Eine womöglich hierarchische Ordnung der Annotationen hinsichtlich ihrer Wiederverwendbarkeit für andere Fragestellungen ist im Detail sicherlich nicht ohne Willkür zu erstellen. Die Annotationen, die den Gesprächsablauf beschreiben und die im Detail oft nur von den bei der Aufnahme Anwesenden beigesteuert werden können, lassen sich aber z.B. gut unterscheiden von den Annotationen, die erst aufgrund der projektspezifischen Analyse notiert werden.

Statt also eine in vielen Fällen willkürliche Unterscheidung in Basis- (oder Roh-) und Fein-Transkript zu treffen, erleichtert es deshalb die Anfertigung und die Weitergabe von Transkriptionen, wenn Annotationen sichtbar und unsichtbar geschaltet bzw. entweder für den Austausch oder Export freigegeben oder aber gesperrt werden können. Um das zu ermöglichen und um gegebenenfalls die Weitergabe von nicht verallgemeinerbaren Annotationen zu unterbinden, müssen *alle Annotationen explizit* als solche *kodiert* und darüber hinaus *typisiert* werden können, damit sich Exportfilter (als Bestandteil eines Transkriptionseditors oder auch getrennt davon) formulieren lassen, die nur genau die Annotationen austauschen, die dafür vorgesehen sind.

1.8. Globale Annotationen ("Transkriptionsköpfe")

Mit den Transkriptionen werden häufig Informationen, wie z.B. Ort und Zeitpunkt der Aufnahme, sprachliche oder biographische Merkmale der beteiligten Sprecher, technische oder sonstige Rahmenbedingungen der Aufnahme dokumentiert. Solche Vermerke beziehen sich auf die Aufnahme und das Transkript als Ganze und nicht nur auf ausgewählte Segmente. Dokumentationen auf der Ebene einzelner Transkriptionen sind *globale Annotationen*.³⁶ Sie sind für die Verwaltung von Transkriptionen in Korpora relevant. Die systematische Dokumentation bzw. Verwaltung von Transkriptionen dient ihrer Auffindbarkeit aufgrund äußerlicher Merkmale und ist eine hilfreiche Voraussetzung für die Nutzung von Retrieval-Systemen oder für die Archivierung.

³⁵ Die Entwicklung von Basis-Transkriptionen zu Fein-Transkripten, also die ständige Weiterbearbeitung ist bei HIAT sogar namensstiftend (ArbeitsTranskriptionen).

³⁶ Je nach Transkriptionskonvention (oder Werkzeug) werden globale Annotationen z.B. als *Transkript-Titel* [HIAT], *Transkriptionskopf* [GAT, DIDA] oder *Meta-Informationen* [EX-MARaLDA] (dazu mehr im Abschnitt 2.2) bezeichnet.

In den verschiedenen Transkriptionskonventionen wird jeweils ein vorsichtiger Versuch unternommen, die zu erfassenden globalen Annotationen zu regeln. Es blieb aber immer nur beim Versuch. Es gibt keinen etablierten Kanon von Dokumentationsinformationen, die immer erhoben werden (sollten). In unterschiedlichen Projekten sind eben sehr unterschiedliche Informationen relevant. Auch ein Minimal-Konsens ist in dieser Frage nicht erkennbar. Das wird dann problematisch, wenn Korpus- oder gar Institutionen-übergreifende Systematiken für die umfassende Recherche nach Transkriptionen aufzubauen sind.³⁷ Die Regelungen, die die Austauschbarkeit einzelner Transkripte sicherstellen können, betreffen in erster Linie das Innere der Transkripte, nicht deren äußere Hülle.

Alle Transkriptionseditoren bieten die Möglichkeit, bei der Neuanlage eines Transkripts globale Annotationen zu erfassen. Diese Angebote werden aber nur selten oder nur unvollständig angenommen; sei es, dass der Zweck der zu notierenden Informationen für das eigene Projekt nicht erkennbar ist oder dass für die im Projekt relevanten Informationen im Transkriptionseditor kein Platz vorgesehen wurde.

Globale Annotationen ließen sich wohl nur zusammen mit Ausführungen zur Korpusverwaltung angemessen behandeln. Das soll hier aber nicht das Thema sein, denn zur Charakterisierung von Transkriptionseditoren ist ihr Beitrag zur Korpusverwaltung nicht wesentlich. Das Transkribieren und das Verwalten des Transkribierten läßt sich nämlich ohne Komplikationen unabhängig voneinander mit getrennten (bestenfalls aufeinander abgestimmten) Werkzeugen bewerkstelligen.³⁸

1.9. Zwischenergebnis

Die hier betrachteten Annotationen zu Transkriptionen werden formal auf vielfältige Weise realisiert. Sie werden überwiegend nicht getrennt vom Transkripttext kodiert, sondern sind in die Texte eingebunden. Keine der Transkriptionskonventionen befaßt sich damit, die Präsentation der Transkripte von der Kodierung der Annotationen zu trennen. Auch die Trennung der Präsentation von den

³⁷ Eine der Initiativen ist die des Open Language Archive Consortiums (OLAC, siehe unter www.OLAC.org). Darin sind diverse Forschungseinrichtungen aus mehreren Ländern zusammengeschlossen, die Internet-Standards für den Zugriff auf Sprachmaterialien vereinheitlichen wollen und Server für diesen Zugriff betreiben.

³⁸ Alle hier erwähnten Transkriptionseditoren bieten rudimentäre Mittel, die der Korpusverwaltung dienen sollen und speichern diese Informationen in den Transkriptdateien oder einer getrennten Datenbank. Es würde jedoch völlig genügen, wenn sich Transkriptionseditoren entweder *gar nicht* um die Verwaltung der Transkripte kümmern müssten oder aber nur insoweit, dass sie sicherstellen, dass ein darauf spezialisiertes System die Transkripte eindeutig identifizieren und mit der andernorts hinterlegten Information verbinden kann. Es scheint bei Gesprächsforschern eine verbreitete, mir jedoch unerklärliche Hochachtung vor derartigen Datenbanken zu geben. Die Programme, die erst einmal "nur" das effiziente Erstellen von Transkripten ermöglichen sollten - das ist kompliziert genug -, wurden meist so konzipiert, dass sie zugleich auch elementare Aufgaben eines Verwaltungssystems für Transkripte wahrnehmen. Siehe DIDA, Transedit (ein bislang unvollendet gebliebener Versuch aus Halle), CLAN, und selbst die in dieser Hinsicht nicht besonders ambitionierten Programme syncWRITER und HIAT-DOS sind da nicht ganz ausgenommen. EXMARaLDA (s.u.) überbietet alle anderen Werkzeuge in diesem Punkt.

anderen Strukturen der Transkripte wird von keiner der Konventionen zwingend gefordert oder auch nur unterstützt.³⁹

Um die Austauschbarkeit von Transkriptionen heute und auch längerfristig universell sicherzustellen und um Textstellen mit Annotaten bestimmten Typs maschinell effizient auffinden, exportieren und auswerten zu können, müssen alle Annotationen explizit kodiert werden, d.h. sind getrennt von ihrer Präsentationsform zu notieren. Zugleich müssen die Werkzeuge zum Transkribieren die oben ausgeführten Strukturen zum Annotieren so unterstützen, dass die Präsentation der Annotate den etablierten Konventionen entspricht.

Die Präsentation der Annotation gleichartiger Phänomene fällt in den verschiedenen Konventionen unterschiedlich aus. Nur ein Transkriptionseditor, der alle Annotationen explizit als solche behandeln und deren Kodierung von deren Präsentation trennen kann, wäre dazu in der Lage, Transkriptionen, die nach den unterschiedlichen Konventionen erstellt wurden, gleichermaßen einzulesen und ineinander zu überführen. Einen solchen Vielzweck-Editor gibt es nicht. Er wird gelegentlich gewünscht⁴⁰ und dem EXMARaLDA-System, das nachfolgend betrachtet wird, eilt der Ruf voraus, dereinst diese Leistung zu vollbringen.⁴¹

2. Zum Konzept von EXMARaLDA

Im ersten Abschnitt wurden aus gängigen Transkriptionskonventionen vielfältige formale Strukturen zum Annotieren von Transkripten hergeleitet. In diesem Abschnitt soll nun das neu entwickelte Transkriptionssystem EXMARaLDA⁴² (EXtensible MARkup Language for Discourse Annotation) betrachtet werden. Es ist eine "Werkzeugkiste" zum Erstellen von Transkriptionen in bestimmten Formaten, definiert aber keine neue Transkriptionskonvention wie CLAN, GAT oder HIAT. Die nachfolgende Darstellung zielt vorrangig darauf zu überprüfen, inwieweit von EXMARaLDA die im ersten Abschnitt herausgearbeiteten Annotationsstrukturen praktisch unterstützt werden.⁴³

EXMARaLDA verwendet zur Strukturierung von Transkripten das Konzept der *Annotationsgraphen*. Der von Bird und Liberman (2001) beschriebene Annotationsgraphen-Formalismus führt sprachliche Strukturen zunächst auf zeitliche Strukturen zurück. Alle Bestandteile eines Transkripts werden aufgefaßt als "Annotationen", die sich zeitlich anordnen lassen. Der Begriff der *annotation* ist dabei ein äußerst allgemeiner und unspezifischer Oberbegriff für alles - meist Schriftliche -, das dem Sprachsignal angeheftet wird bzw. werden könnte. Auch die Verschriftlichung des Gesprochenen selbst besteht im Konzept der Annotationsgra-

³⁹ Von der CLAN-Konvention wird diese Trennung sogar ausgeschlossen.

⁴⁰ Ob dieser Wunsch einem wirklichen praktischen Bedarf entspricht, oder ob er nicht eher eine Art Sehnsucht nach Frieden unter den Kollegen widerspiegelt, die sie sich mit der Auswahl einer Konvention immer zugleich gegen eine andere entscheiden müssen, kann ich nicht entscheiden.

⁴¹ Siehe Dittmar (2002:179). Die Antwort darauf ist zu finden bei Schneider (2002:38-43).

⁴² Zur Konzeption und Implementierung von EXMARaLDA siehe Schmidt (2002a und 2002b).

⁴³ Insbesondere soll dieser Text kein Testbericht sein und will auch keinem Interessenten die Mühe und das Vergnügen (ab)nehmen, den EXMARaLDA-Editor selber zu erkunden. Die Dokumentation und das Editor-Programm ist aus Hamburg vom SFB 538 zu haben unter www.rrz.uni-hamburg.de/exmaralda. Für diesen Text wurde zuletzt die Version 1.2.2 (2002-11-29) betrachtet.

phen aus "Annotationen". Ich dagegen verwende hier den Begriff in einem wesentlich engeren Sinne: zu den Annotationen zähle ich - kurz gesagt - nur die "Ergänzungen" des verschriftlichten Gesprochenen, des Transkripttextes.⁴⁴

Das theoretische Konzept der Annotationsgraphen ist offen für unterschiedliche Implementierungen, EXMARaLDA ist nur eine davon.

2.1. DV-technische Merkmale

EXMARaLDA ist eine Sammlung von formalen, also nicht an den Inhalt gebundenen, DV-technischen Konventionen zur Kodierung von Transkripten und dazu passenden Werkzeugen. Zu den Werkzeugen gehören Konvertierprogramme, die Transkripte importieren, aus den verschiedensten EXMARaLDA-Formaten in andere interne Formate konvertieren und auch in einige externe, vor allem zur Präsentation und Nachbearbeitung geeignete Datenformate exportieren können. Als zentrales Werkzeug fungiert ein Transkriptionseditor, mit dem man Transkripte im *basic-transcription*-Format erzeugen kann. Dieses Werkzeug dürfte wohl der für die meisten Anwender sichtbarste Bestandteil des EXMARaLDA-Systems sein.

2.1.1. EXMARaLDAs DV-Technik

Die Programme und Daten wurden mit modernen Techniken erstellt: Programme in Java und die Definition der Datenstrukturen in XML (= eXtensible Markup Language). Mit Java und XML werden systemseitig automatisch auch Unicode-Zeichen verfügbar.⁴⁵

Die Programme können "im Prinzip" auf allen Systemen ablaufen, für die eine funktionierende Java-Laufzeitumgebung verfügbar ist, das sind zumindest alle relevanten Betriebssysteme für Arbeitsplatz-Computer: Windows (ab Win'98), OS/X für Macs und Linux. Gelegentlich kämpft der EXMARaLDA-Editor noch mit kleinen Problemen, die in Java-Bibliotheken oder dem Laufzeitsystem begründet sind.⁴⁶

⁴⁴ In gewissem Sinne verhalten sich Annotationen (in meinem Sinne) zum Transkripttext wie die "Paratexte" im Sinne von Genette (2001) zum literarischen Text. Genette stellt das textuelle "Beiwerk" zur "eigentlichen" Literatur sehr formal dar, erklärt dessen Funktionen und belegt seine Charakterisierungen mit konkreten Beispielen aus seinem Korpus. Seine Methode ist der des ersten Teils dieses Artikels verwandt, wenn er auch ein literarisches und zudem sehr viel umfangreicheres Korpus verschiedener "Textarten" behandelt. Seine detaillierten Ausführungen zu den "Anmerkungen" (S.304-327) als einer eigenständigen "Gattung" von Paratexten und ihres Verhältnisses zum Text treffen übrigens vielfach auch auf Annotationen in Transkripten zu.

⁴⁵ Dass diese in allen modernen Betriebssystemen schon seit längerem intern enthaltene Möglichkeit von den EXMARaLDA-Programmen und -Datenstrukturen auch so genutzt wird, dass Anwender davon profitieren können (sichtbar wird dies z.B. bei den IPA-Zeichen), ist bemerkenswert, weil heute leider immer noch nicht selbstverständlich.

⁴⁶ Ein Vorteil von Java-Programmen, dass sie nämlich dem Programmierer wenig Gelegenheit bieten, tückische, in anderen Programmiersprachen nur mühsam vermeidbare Speicher-Verwaltungsfehler einzubauen, die eventuell sogar den ganzen Rechner lahmlegen können, wird mit einem hohen Ressourcenbedarf erkaufte, d.h. Java-Programme laufen in der Regel stabil, erfordern aber auch leistungsfähige Rechner.

Wichtiger noch als die Austauschbarkeit der Programmumgebung dürfte die der Transkripte sein. Mit der Verwendung von Datenstrukturen, die mit Mitteln von XML definiert wurden, ist Austauschbarkeit der Transkripte "im Prinzip" gewährleistet. XML ist im doppelten Sinne ein "offenes Format": Die Definition von XML erfolgte beim W3C (WWW-Consortium) und steht jedem kostenlos zur Verfügung; und XML-formatige Dateien sind für jeden lesbar, da sie reine Text-Dateien und zudem weitgehend selbsterklärend sind. Die Daten können notfalls ohne besondere Hilfsmittel gelesen und geschrieben werden. Für einen effizienten Einsatz der Daten ist natürlich ein dafür optimiertes Werkzeug erforderlich.

Im EXMARaLDA-System sind vier Transkript-DTDs definiert. DTD steht für *Document Type Definition*. Eine Transkript-DTD ist hier nichts anderes als die Definition der abstrakten Strukturen der Textart "Transkript" mit formalen (DV-technischen) Mitteln gemäß den Regeln der XML. Es sind

- die *list-transcription.dtd*, die für Transkriptionen in einem Listenformat (z.B. im CHAT-Format) vorgesehen ist,
- die *basic-transcription.dtd*, die für Transkriptionen in einem Partiturformat, das vom EXMARaLDA-Editor gelesen und geschrieben werden kann, verwendet wird,
- die *segmented-transcription.dtd*, die das zentrale Format mit den umfangreichsten Notationsmöglichkeiten darstellt (alle EXMARaLDA-Transkripte können in dieses Format gebracht werden und alle EXMARaLDA-Formate lassen sich daraus gewinnen - teilweise jedoch nur mit konzeptionell bedingten Informationsverlusten),
- die *interlinear-text.dtd*, die ein Format zur Erstellung von Präsentationsformaten (z.B. HTML oder RTF) für Transkripte im Partiturformat definiert.

2.1.2. Export von Transkripten⁴⁷

Für den *Export* von Transkripten werden von EXMARaLDA derzeit mehrere Formate berücksichtigt: für die Präsentation von Transkripten RTF (für den Übergang zur Word-Textverarbeitung) und HTML, für die weitere Analyse TASX,⁴⁸ textgrids des praat-Systems und AIF für das AGTK.⁴⁹ Ein Export in das CHAT-Format ist geplant. Die Liste der Export-Formate läßt sich aufgrund des modularen Konzepts von EXMARaLDA bei Bedarf erweitern.

⁴⁷ Der Import von Transkripten wird erst nach der Darstellung der Annotationen in EXMARaLDA erläutert, siehe unter 2.6.

⁴⁸ TASX (Time Aligned Signal data eXchange) ist ein System, das neue Technologien für den Zugriff auf Korpusmaterialien über Netzwerke bereitstellt, darunter ein einfacher Transkriptionseditor, im TASX-System Annotator genannt. Siehe auch Fn 71 und unter tasxforce.lili.uni-bielefeld.de

⁴⁹ AGTK (Annotation Graph Toolkit) ist eine Implementierung von Datenstrukturen nach dem Annotationsgraphen-Konzept. Das AGTK-Projekt wird vom Linguistic Data Consortium (LDC an der University of Pennsylvania) betrieben, um den Austausch von Daten in diversen Annotationsformaten zu vereinfachen, siehe unter agtk.sourceforge.net.

2.1.3. Exemplarische EXMARaLDA-Daten

Das oben unter 1.1 gezeigte fiktive Beispiel könnte im EXMARaLDA-Editor (transkribiert "in Anlehnung an" HIAT, d.h. soweit sich HIAT-Konventionen mit EXMARaLDA realisieren lassen) so aussehen:⁵⁰

	0 [240.]	1 [24.]	2 [244.]	3 [244.]	4 [245.]	5 [245.]	6 [247.]	7 [247.]	8 [248.]
S1 [v]	Kommst du morgen mit uns mit?			zum Sch	wimmen.				
S1 [nv]									
S2 [v]		hm	((0,5s))		mór	gen lau/ äh	fáhr	ich	in die
S2 [i]						----			
S2 [nv]						zeigt mit dem Daumen zum Fenster			
[wav]	Tonaufnahme								

Der EXMARaLDA-Export (via RTF) führt zu folgender Darstellung:⁵¹

S1 [v]	Kommst du morgen mit uns mit?	zum Schwimmen.
S2 [v]	hm ((0,5s))	mór gen
[wav]	Tonaufnahme	

S2 [v]	lau/ äh fáhr ich	in die Stadt.
S2 [i]	----	
S2 [nv]	zeigt mit dem Daumen zum Fenster	
[wav]		

Im XML-Format gemäß der *basic-transcription.dtd*, das ist EXMARaLDA's internes Format für den Partitur-Editor, sähe das auszugswise⁵² etwa so aus:

```
<tier id="TIE1" speaker="SPK1" category="v" type="t">
<event start="T0" end="T1">Kommst du morgen mit uns
mit?</event>
<event start="T3" end="T4">zum Sch</event>
<event start="T4" end="T5">wimmen.</event>
</tier>
<tier id="TIE2" speaker="SPK1" category="nv" type="a">
</tier>
<tier id="TIE0" speaker="SPK2" category="v" type="t">
<event start="T1" end="T2">hm` </event>
<event start="T2" end="T3"> ((0,5s)) </event>
<event start="T4" end="T5">mór</event>
<event start="T5" end="T6">gen lau/ äh </event>
<event start="T6" end="T7">fáhr </event>
<event start="T7" end="T8">ich </event>
<event start="T8" end="T9">in die Stadt. </event>
```

⁵⁰ Insbesondere die HIAT-konforme Notation der Intonation könnte in EXMARaLDA auf verschiedene Weise erfolgen: wie hier intralinear mit Diakritika oder oberhalb in einer getrennten Zeile mit einer Synchronisation der Position des Intonationssymbols.

⁵¹ Da beim RTF-Export derzeit noch leichte Fehler bei der Berechnung der Laufweite von Schriften auftreten, die sich auf den Umbruch der Tabellenzellen auswirken, wurde zur sauberen Darstellung der nachfolgende Partiturblock manuell etwas nachgebessert.

⁵² Die komplette EXMARaLDA-XML-Datei dieses Demo-Transkripts im basic-transcription.dtd-Format steht - zusammen mit kommentierenden Lesehilfen - im Anhang.

```

</tier>
<tier id="TIE3" speaker="SPK2" category="i" type="a">
<event start="T6" end="T7">---- </event>
</tier>
<tier id="TIE4" speaker="SPK2" category="nv" type="a">
<event start="T6" end="T8">zeigt mit dem Daumen zum Fenster
</event>
</tier>

```

Wie das Beispiel andeutungsweise erkennen läßt, legt EXMARaLDA konzeptionell Wert darauf, die *äußere Darstellung von der internen Speicherung der Transkripte zu trennen*. Das ist im Prinzip die richtige Vorgehensweise, sie wurde bisher nur selten praktiziert.⁵³ Leider geht EXMARaLDA bei dieser Trennung jedoch nicht konsequent genug vor. Wie in EXMARaLDA transkribiert und annotiert werden kann, wird in den folgenden Abschnitten im Detail erläutert.

2.2. Globale Annotationen in EXMARaLDA

Die Darstellung der Annotate soll mit den globalen Annotationen⁵⁴ beginnen, weil mit ihnen auch das Gerüst definiert wird, in dem alle anderen Annotationen ihren Platz finden müssen. Ohne das Verständnis der übergeordneten Strukturen lassen sich die Möglichkeiten und Grenzen zur Strukturierung von Annotationen mit geringerer Reichweite nicht verstehen.

Korpusverwaltung und Sprecher-Merkmale

Im EXMARaLDA-System wurde für jedes Transkript ein flexibler Bereich mit Meta-Informationen vorgesehen. Der (für alle EXMARaLDA-DTDs identische) "head" enthält Felder zur Kennzeichnung des Projektes (oder Korpus), des Namens der Transkription und einer URL, die mit dem Transkript verbunden ist, und der eingesetzten Transkriptionskonvention. Es lassen sich außerdem vom Benutzer eigene Meta-Informationen definieren, d.h. neue Felder erzeugen, mit denen die Transkription beschrieben werden kann.

Neben der ganzen Transkription lassen sich einzelne Sprecher charakterisieren. Es sind Felder für Sprecher-Sigle und -Name, Geschlecht, die verwendeten Sprachen und die Erst- und Zweitsprache (L1 und L2) vorbereitet. Solche Sprecherinformationen sind gewöhnungsbedürftig. Die in den DTDs verankerten Sprachen-Felder⁵⁵ entsprechen einem speziellen Bedarf am SFB 538. In anderen Zusammenhängen dürfte die Notation von L1 und L2 irrelevant sein. Es können auch auf der Ebene jedes Sprechers - wie schon auf der übergeordneten Ebene -

⁵³ Zu den Ausnahmen gehören die in SGML-kodierten Transkripte der TEI (Text-Encoding-Initiative), Siehe dazu Sperberg-McQueen/Burnard (1994) oder zum aktuellen Stand der Initiative unter www.tei-c.org, darunter besonders das Kapitel 11 "Transcriptions of Speech".

⁵⁴ Ein Teil der globalen Annotationen wird im EXMARaLDA-Editor unter dem Titel "Meta-Information" verwaltet, ein anderer in einer "speakertable" mit "speaker properties" und ein weiterer heißt "tier properties".

⁵⁵ Die Sprachen-Felder dürfen zwar leer bleiben, sie tauchen aber sogar bei der Präsentation der Transkripte wieder auf.

zusätzliche eigene Felder definiert werden. Mit diesem Verfahren wird eine erfreulich hohe Flexibilität der Dokumentation erreicht.⁵⁶

Partitur-Strukturen

Wie an obigem Beispiel ersichtlich, sind EXMARaLDA-Transkripte aus Partiturzeilen (tiers) aufgebaut, die wiederum in Zellen zerlegt werden. Den Zerlegungspunkten können Zeitangaben hinzugefügt werden. Die Zerlegung der Zeilen in Tabellenzellen kann über mehrere Zeilen hinweg erfolgen, muss aber nicht alle Zeilen betreffen. EXMARaLDA-Transkripte bestehen also aus Tabellen-Strukturen mit durchgängig definierten Zeilen und nicht durchgängig definierten Spalten.⁵⁷

Partiturzeilen werden unabhängig voneinander definiert. Diejenigen Zeilen, die mit einer identischen Sprecher-ID gekennzeichnet wurden, sind lose über die Sprecherzugehörigkeit miteinander verbunden. Ihre Reihenfolge ist völlig frei. Sie lassen sich nachträglich vertikal umordnen oder (un-)sichtbar schalten. Es lassen sich sogar Sprecherdefinitionen löschen, ohne dass damit auch die zugehörigen Zeilen verloren gehen. Es gibt im EXMARaLDA-System also - anders als bei IDS-DIDA und HIAT-DOS - keine expliziten Sprecherblock-Strukturen. Sprecherblöcke gehören zu genau einem Sprecher, sind also Teilmengen der Partitur. Sie fassen alle Zeilen eines Sprechers zusammen, sind also Obermengen für Zeilen mit unterschiedlichen Funktionen. Die Zeilen eines Sprecherblocks sollten gemeinsam definiert und bearbeitet werden können. In EXMARaLDA müssen die Transkribenten die Zusammengehörigkeit von Partiturzeilen zu Sprecherblöcken aber manuell simulieren, da sie im technischen Sinne keine strukturierte Einheit darstellen, allenfalls lose Sammlungen.

Strukturvorgaben

Die Definition der Felder für globale Annotationen basiert nicht auf Vorlagen. Für jedes Transkript ist immer wieder neu manuell sicherzustellen, dass alle benötigten eigenen Felder definiert und die erforderlichen Zeilen für jeden Sprecher angelegt werden.⁵⁸ Es gibt nur für die Definition der verwendeten Schriftarten die Möglichkeit, sie abzuspeichern und später zu den Transkripten wieder hinzuzuladen.⁵⁹

⁵⁶ Man könnte also z.B. die Felder, die die "Deckblätter" des DIDA-Editors oder HIAT-DOS zur Kennzeichnung der Transkripte bereithalten, oder die Transkriptionsköpfe des CHAT-Systems auch im EXMARaLDA-Editor nachbilden.

⁵⁷ Die Strukturen anderer Partitureditoren sind weniger tabellenförmig: IDS-DIDA und HIAT-DOS kennen keine Zerlegung der Zeilen in Spalten und im syncWRITER-Programm lassen sich mit synctabs zwar auch Zeitpunkte definieren, aber die zerlegen die Zeilen nicht in unterschiedlich viele Zellen.

⁵⁸ Eventuell sollte man sich für ein Projekt ein leeres Transkript mit vielen (leeren) Sprechern als Vorlage erzeugen, in dem die im Projekt benötigten Felder und Zeilen schon definiert sind. Es ist zuverlässiger und schneller, diese Vorgaben und das leere Gerüst nachträglich individuell abzuwandeln (bei weniger Sprechern die überzähligen Zeilen auch wieder zu löschen) als die Felder und Zeilen jedesmal neu anzulegen.

⁵⁹ Diese Definitionen sind jedoch nur dann auf andere Transkripte übertragbar, wenn der Aufbau der Partituren sehr ähnlich ist und zudem deren interne Kennungen (IDs) sehr genau geplant werden.

Verwendbarkeit der globalen Annotation

Im EXMARaLDA-Editor lassen sich zwar vielfältige globale Annotationen vornehmen, für die Korpusverwaltung kann man sie innerhalb dieses Transkriptionseditors aber kaum verwenden. Für ein effizientes Retrieval in Korpora ist ein weiteres Werkzeug, das die Transkriptionsköpfe auslesen und effizient datenbankgemäß behandeln kann, erforderlich. Das ist zukünftig noch zu entwickeln.⁶⁰ Transkripte benötigen eigentlich überhaupt keine "head"-Struktur (zumindest keine so umfangreiche). Sie müssen die Informationen, die für ihre Verwaltung sinnvoll sind, nicht selbst enthalten. Die Korpusverwaltung erfolgt ohnehin getrennt vom Editor. Als Bindeglieder zwischen einem Werkzeug zur Korpusverwaltung und den Transkripten würden eindeutige Kennungen (z.B. als Teil der Namen) für Transkript-Dateien genügen.

Die derzeit in EXMARaLDA zur Notation globaler Annotationen intern verwendeten XML-Strukturen bieten flexible Möglichkeiten für eigene Meta-Informationen zur Verwaltung der Transkriptionen. Diese Annotationen sind zwar offen für die Weiterverarbeitung mit anderen Programmen, sie wurden intern jedoch leider so angelegt, dass dies mit einigen technischen Umständlichkeiten verbunden ist.⁶¹

2.3. Strukturtypen für Annotate in EXMARaLDA

In diesem Abschnitt wird dargestellt, wie sich die im Abschnitt 1.3 herausgearbeiteten Strukturtypen A bis D für Annotate mit EXMARaLDA realisieren lassen.

Im EXMARaLDA-Editor und in den DTDs ist für Segmente ausschließlich die Struktur *Ereignis* ("event" von T_i bis T_j), vorgesehen. Segmente werden im Editor als Zellen in Zeilen gehandhabt. Der Editor unterstützt nur Transkripte, die

⁶⁰ Eine "echte" Datenbank zur Korpusverwaltung ist für den Bedarf des SFB 538 wohl auch geplant.

⁶¹ Die EXMARaLDA-internen XML-Strukturen zur Notation der globalen Annotationen sind uneinheitlich. Teilweise werden die Informationen in Attribute verlegt, teilweise sind sie als Texteinträge zwischen Tags angeordnet. z.B. wird `<transcription-name>Interview 123</transcription-name>` verwendet. Das ist für notwendige Informationen in Ordnung. Aber warum gibt es daneben eine Struktur, in der die Information in ein Attribut verlagert wurde: `<referenced-file url="IV123.wav"/>` und nicht strukturell so wie oben, also: `<referenced-file-url>IV123.wav</referenced-file-url>`? Dasselbe noch einmal an einem anderen Beispiel: `<sex value="f"/>`, alternativ könnte man diese Information auch mit der Struktur `<sex>f</sex>` notieren (dann aber nicht mehr automatisch auf richtige Einträge prüfen lassen). Um die globalen Annotationen in einheitlicher Struktur zu deklarieren, wäre die erweiterbare Struktur, die schon für die vom Anwender definierbaren Informationen zu Einsatz kommt, zu bevorzugen. Auf die Kennzeichnung "vom Benutzer definiert" könnte man verzichten (oder in einem optionalen Typ-Attribut notieren). Obiges Beispiel ließe sich dann (wenn es denn schon in englischer Sprache sein soll) wie folgt kodieren: `<info name="sex">f</info>`

Durch eine Vereinheitlichung der heterogenen Strukturen würde die Weiterverarbeitung (z.B. für ein Retrieval) wesentlich einfacher. Der Nachteil der Verlagerung von Informationen in den Textbereich zwischen `info`-Tags besteht darin, dass sich mit üblichen XML-Validatoren nicht mehr prüfen läßt, ob eine bestimmte Information vorhanden ist oder vorgesehene Werte enthält (z.B. `sex=m|f|u`). Dieser Mangel besteht aber bei allen vom Benutzer definierten Informationen auch jetzt schon. Die Prüfung der Daten auf Vollständigkeit und Konsistenz kann deshalb nicht nur mit einem einfachen XML-Parser erfolgen. Hier wäre z.B. eine Vorlagen-Technik, die von EXMARaLDA aber leider nicht verwendet wird, ein sinnvolles Verfahren.

gemäß der *basic-transcription.dtd* gestaltet sind.⁶² Es gibt keine direkte Möglichkeit den einzelnen Segmenten weitere Eigenschaften zuzuordnen. Das geht nur für ganze Zeilen. Sie lassen sich mit zwei unabhängigen Eigenschaften versehen: "category" und "type". Da die *Ereignis*-Strukturen keine anderen Eigenschaften außer Anfangs- und Endezeitpunkt haben, kennzeichnen sie nur *Zeiträume*.

2.3.1. Annotierende Textattribute (Strukturtyp A)

Annotierende Textattribute lassen sich mit EXMARaLDA kaum realisieren. Annotationen, die mit Versalien notiert werden ("`lAUter`"), kann man im Editor zwar wie gewohnt notieren. Diese Annotation wird aber in der XML-Struktur nicht explizit als solche gekennzeichnet und ist deshalb XML-technisch wertlos. Hier bliebe wie bisher die Präsentation mit der Kodierung vermischt. Annotationen, die mit einer anderen Änderung der Schriftauszeichnung zu notieren sind ("`lauter`"), lassen sich mit EXMARaLDA ebenfalls nicht wie gewohnt realisieren, weil sich in EXMARaLDA die Auswahl einer Schriftart oder Schriftauszeichnung immer auf ganze Zeilen bezieht und nicht auf einzelne Zellen in einer Zeile.⁶³ Die Annotation durch Unterstreichung lässt sich, wie obiges Beispiel unter 2.1 zeigt, durch Einfügen einer Annotationszeile unterhalb der Zeile für die Notation des Gesprochenen allenfalls schlecht imitieren,⁶⁴ bei der Annotation durch Fettschrift oder gar durch einen Wechsel der Schriftart ginge auch das nicht. Die Präsentation des Transkripts wird in solchen Fällen nicht den Transkriptionskonventionen entsprechen. Als Ausweg könnte man erwägen, prinzipiell eine zweite Zeile für Annotationen des Strukturtyps A zu definieren, in die man die Annotate für den in der darüber liegenden Zelle stehenden Text ("`au`", aus "`l|au|ter`", das mit zwei Zeitmarken auf drei Zellen verteilt wird) mit platzsparenden Kürzeln einträgt (z.B. "`bB`" für "besondere Betonung"). Das sähe dann in etwa so aus:

```
S1_Text: ich spreche l|au|ter als sonst
S1_Anno:                |bB|
```

⁶² Zitat aus der EXMARaLDA-Dokumentation unter "Systemarchitektur und DTDs": "Es ist weiterhin vorgesehen, daß nachträgliche Annotationen (d.h. Annotationen, die nicht während des ursprünglichen Transkriptionsprozesses vorgenommen werden, insbesondere sogenannte 'Stand-Off-Annotations') sich ebenfalls lediglich auf eine segmentierte Transkription beziehen. In der vorliegenden Version 1.0 spezifiziert EXMARaLDA allerdings weder, wie solche Annotationen in XML kodiert werden, noch wie sie zusammen mit der Transkription in ein darstellungsnäheres Format wie das der Basis-, Listen- oder Partitur-Tabellen-Transkription konvertiert werden können. Dies wird in einer späteren Version nachgeholt werden."

⁶³ Die Begründung ("Beispielsweise ist es nicht möglich, Unterstreichungen oder Fettdruck in eine Datenbank zu übernehmen." siehe EXMARaLDA-Handbuch, Version 1.2.2, p.15) überzeugt nicht. Hier wird das "Kind mit dem Bade ausgeschüttet" statt es zu (be)reinigen: Die angestrebte Trennung von interner Kodierung und äußerlicher Präsentation der Annotationen führt dazu, dass Annotationen durch Schriftauszeichnungen wie z.B. Unterstreichungen innerhalb des EXMARaLDA-Systems überhaupt nicht mehr vorgesehen sind. Aus der sinnvollen Trennung wird ein ein falscher Gegensatz: die Kodierung verdrängt hier die Präsentationsform.

⁶⁴ Die Imitation ist unbefriedigend: Der vertikale Abstand zwischen dem Text und der "Unterstreichung" bleibt zu groß. Ferner muss man die Länge des zu annotierenden Textes in der Annotation selber berücksichtigen. Das wird bei proportionaler Schrift nicht immer gelingen. Und es ist vor allem - trotz der Verlagerung der Annotation in eine getrennte Zeile - keine wirkliche Trennung von Kodierung und Präsentation der Annotation.

Systemtechnisch sind die beiden Spuren, die man hier benötigt (hier: `S1_Text` und `S1_Anno`), nur sehr lose aneinander gebunden. Dass es sich beim Inhalt der zweiten Spur um die zur ersten Spur gehörenden Annotationen handelt, ist in den Datenstrukturen von EXMARaLDA nicht zwingend festgelegt. Nur die Sorgfalt der Transkribenten und die Interpretation verständiger Leser stellt sicher, wie die Zeilen gemeint sind.⁶⁵ Man könnte das zwar mit gesonderten Regeln für die Verwendung der Sprecher-Siglen und für die Positionierung der Spuren zueinander ("oben den Text, darunter die Annotate") fixieren. Sie sind systemtechnisch mit den aktuellen Versionen der zugrunde liegenden EXMARaLDA-DTDs aber nicht definierbar und deshalb mit XML-Mitteln auch nicht validierbar. Der Austausch solcher Informationen bleibt deshalb heikel.

In einem späteren (manuellen?) Verarbeitungsschritt müsste man dann in die vom Editor erzeugte Präsentation (z.B. im RTF-Format) mit einem anderen Textverarbeitungsprogramm an den annotierten Stellen die Textattribute einarbeiten und anschließend die Annotationszeile aus der Präsentation tilgen.

Effizienter wäre es, die zu annotierenden Textsequenzen zu markieren und den Markierungen ein Attribut (z.B. "besondere Betonung") zuweisen zu können, das mit expliziten XML-Strukturen als Annotat (mit Angaben zu seiner Reichweite) festgehalten wird. Wie die Annotation dann am Bildschirm und im Ausdruck präsentiert wird, könnte außerhalb des Transkripts in einer Layout-Beschreibung, die zu der verwendeten Transkriptionskonvention passen muss, festgelegt werden. Bei diesem Beispiel könnte dann statt "`lauter`" für eine andere Transkriptionskonvention auch "`lAUter`" ausgegeben werden. DV-technisch ermöglicht streng genommen erst die *konsequente* Trennung von Annotationen und Transkripttext die effiziente Weiterverarbeitung, z.B. die Suche nach bestimmten Typen von Annotationen (hier "besondere Betonung") und nach den Wörtern (hier nach dem Wort "lauter").

2.3.2. Intralineare Annotationen (Strukturtyp B)

Intralineare Annotationen, d.h. punktuelle Einfügungen von kurzen Texten oder Zeichen in den Transkripttext ("`((3, 5s))`", "`la: :ngsa:m`"), lassen sich im EXMARaLDA-Editor wie in jedem Texteditor notieren. EXMARaLDA behandelt diese Einschübe aber nicht als Annotate, d.h. sie werden in den XML-Strukturen des Transkripts nicht explizit als solche ausgewiesen. Damit sind sie technisch ebenso bedeutungslos und wertlos wie die oben schon angeführten annotierenden Versalien.

Eine allgemeine Empfehlung, wie man als Transkribent mit diesem Manko umgehen sollte, lässt sich kaum geben. Da es innerhalb von Zeilen nur Zellen mit Angaben zur Anfangszeit und Endezeit gibt, wären auch hier weitere Zeilen erforderlich, um parallel zu den Zellen, deren Inhalt annotiert werden soll, XML-Strukturen zu verankern, in denen die Annotate festgehalten werden können. Die bisherige Form bliebe im Transkripttext erhalten und würde um explizite Annotate in einer zweiten Zeile ergänzt. Für Präsentationszwecke müsste man die Ver-

⁶⁵ XML-technisch sind Bindungen von Zeilen an Sprecher (über eine *speaker-id*) vorgesehen, die Verwendung der Zeilen-type und category ist jedoch recht frei und wie die Sigle ein funktionsloses Etikett.

doppelung der Annotate durch Ausblenden der zweiten Zeile wieder unsichtbar machen, oder die Präsentation der Transkripte würde wesentlich dadurch verändert, dass Annotationen nur noch in einer zweiten Zeile vorgenommen würden.

Die erste Variante sähe dann in etwa so aus (die Annotationszeile und die Zellengrenzen könnten in der Präsentation ausgeblendet werden):

```
S1_Text: ich spreche l|a::|ngs|a:|mer als sonst
S1_Anno:          |sDe|   |De|
```

Die zweite Variante z.B. so:

```
S1_Text: ich spreche l|a |ngs|a |mer als sonst
S1_Anno:          |sDe|   |De|
```

(mit den Kürzeln "De" für Dehnung und "sDe" für starke Dehnung.)

Beides sind keine befriedigenden Lösungen. Man kann, auch wenn dies aus Platzgründen wünschenswert wäre, nicht auf die Kennzeichnung der Segmentenden in der Textzeile verzichten, da andernfalls die Reichweite der Annotation fehlerhaft wäre.

2.3.3. Annotationen von Segmenten in parallelen Zeilen (Strukturtyp C)

Annotationen von Segmenten und deren Extension in parallelen Zeilen sind mit EXMARaLDA gut zu realisieren, da diese Annotationsstruktur in den EXMARaLDA-Datenstrukturen (als einzige) vorgesehen ist. Die Kennzeichnung der Reichweite der Annotation in der Präsentation durch Auffüllen mit Bindestrichen oder der Anfangs- und Endezeitpunkte durch öffnende und schließende Klammern ist in EXMARaLDA nicht so einfach umsetzbar, da man dazu die genaue Laufweite der Schriften beachten muss und der EXMARaLDA-Umbruch diese Zeichenketten ebenfalls an den richtigen Stellen zerlegen müsste (was er verständlicherweise nicht tut). Solche vom Layout abhängigen Kennzeichnungen sollten nicht mit den eigentlichen Annotationen vermischt werden. Man könnte solche Darstellungsformen in der fertigen Präsentationsversion eines Transkripts außerhalb von EXMARaLDA jedoch schnell ergänzen, wenn die Grenzen der Zellen mit den Annotationen erst einmal definiert sind.

2.3.4. Annotation mit Anfangs- und Ende-Symbolen (Strukturtyp D)

Die Annotation durch Einfügung von Symbolen zur Kennzeichnung des Anfangs und Endes eines zu annotierenden Segmentes läßt sich in EXMARaLDA strukturell recht gut erfassen. Das Layout der Präsentation wird sich jedoch nicht immer in gewohnter Weise erstellen lassen. Bei der Kennzeichnung von Simultanpassagen dürfte es kaum Probleme geben. Die in einigen Transkriptionskonventionen vorgesehene Kennzeichnung synchroner Zeitpunkte ist nicht so einfach zu haben, z.B nach IDS-DIDA:

```
S1: ja |aber   |
S2:   |nein nie|mals
```

Die explizite Kennzeichnung von Simultanpassagen durch Sonderzeichen (im IDS-DIDA ist dafür der senkrechte Strich "|" vorgesehen) könnte in EXMARaLDA entfallen, da diese Kennzeichnung schon mit der Struktur der Zellen in den Zeilen sichergestellt ist. Die Grenzen der Zellen können mit Gitterlinien angezeigt werden. Um dies für den Leser noch deutlicher werden zu lassen, könnte man solche Kennzeichen nachträglich in der Präsentationsform ergänzen, ohne dass dadurch die Struktur des eigentlichen Transkripts tangiert wäre.

Um verschachtelte Annotationen von sich überlappenden Segmenten vornehmen zu können, wie sie im Beispiel nach den Konventionen des FKO vorkommen

S2: k+ (ja6) +k

S1: i+ k+ den z+ Goethe +z +k nicht zu zitieren6 +i

sind in EXMARaLDA mehrere Zeilen pro Sprecher erforderlich. Die obige Darstellung aus dem FKO-Korpus mit den abgrenzenden Symbolen läßt sich nicht sinnvoll in einer Zeile pro Sprecher nachbilden, wenn diese Annotationen zugleich in XML-Strukturen abgebildet werden sollen.

Die in Transkriptionen nach GAT gelegentlich verwendeten Klammerungen für längere Kommentare (siehe 1.1.3) lassen sich in einer Kommentarzeile unterbringen, und die Reichweite des Kommentars ist eindeutig mit der Synchronisierung der Zellen zu kennzeichnen. Bei der Präsentation weicht die Darstellung aber zwangsläufig von den GAT-Gewohnheiten ab.

Auch bei Annotationen und Kommentaren, die bisher in einer weiteren Spalte oder in Fußnoten untergebracht wurden, läßt sich mit EXMARaLDA die Reichweite durch die Ausdehnung der Zelle, in der sie notiert sind, sehr gut kennzeichnen. Werden die Kommentare direkt in die Zelle eingetragen, wird ihnen auch in der Präsentation nur eine Zeile eingeräumt. Je nach Länge des Kommentars und des Kommentierten wird letztere eventuell gestreckt. Ein Umbruch des Kommentars in mehrere Zeilen innerhalb der Zelle ist nicht vorgesehen und könnte nur - mit bestimmten Problemen bei der Sicherstellung der Synchronität - nachträglich manuell in der Präsentationsversion erfolgen. EXMARaLDA bietet für längere Ergänzungen auch die Möglichkeit, in den Zellen einer Zeile mit geeignetem Typ Links zu notieren. Das ist vor allem dazu gedacht, Bilder, Töne oder Filme mit dem Transkript zu verknüpfen. Es läßt sich aber auch für eine Verknüpfung mit einer externen Datei, die einen längeren Kommentar enthält, nutzen. Den Kommentar bei der Präsentation passend in das Transkript einzublenden liegt derzeit aber außerhalb der Möglichkeiten von EXMARaLDA.

2.4. Intonation

Bei allen Werkzeugen zur Transkription gemäß den HIAT-Konventionen besteht bisher der Mangel, dass die Kennzeichnung der Intonation zwar mit Symbolen (ähnlich wie Diakritika) erfolgt - und insofern eigentlich vom verschriftlichten Gesprochenen, dem "Text", getrennt ist -, aber systemtechnisch wie Layout behandelt wird. Dieser Mangel, Intonation durch Layout zum Ausdruck zu bringen, wird mit EXMARaLDA nicht überwunden. Es gibt keine Strukturen zur Notation von Intonation. Zwar lassen sich mit Unicode-Mitteln Diakritika über Grundbuch-

staben schreiben (z.B. "m̂ " ⁶⁶), aber die Notation der Intonation wird entweder, wie schon im syncWRITER, mit dem Text, dessen Intonation zu notieren ist, verschmolzen (verändert also den Text) oder wird in einer darüber liegenden Zeile notiert und etwas umständlich durch Einfügung von Synchronankern mit dem zugehörigen Text verbunden. Die bei der Erfassung sehr effiziente Technik, die im HIAT-DOS-Programm angewendet wird, nämlich mit Leerzeichen und Sonderzeichen in einer Zeile über dem Text an den richtigen Stellen die Intonationssymbole zu positionieren, ließe sich nur mit Mühe imitieren und ergäbe auch keine strukturierte Annotation der Intonation, sondern orientiert sich einzig an deren Darstellungsform.

2.5. Volltext-Recherche, Lemmatisierung, Zerlegungen und Bindungen

Die Art der Implementierung der Annotationsgraphen führt bei EXMARaLDA dazu, dass bei Überlappungen Worte in verschiedene Zellen zerfallen. Bei einer präzisen Annotation verteilt der Editor Wörter in mehrere Zellen. Aufgrund der Zerlegungen durch Zeitanker in *events* sind die dann mit einfachen Suchwerkzeugen in den Transkripten gemäß der *basic-transcription.dtd* nicht mehr auffindbar. In Transkriptionen gemäß der *segmented-transcription.dtd* (die kann der Editor nicht direkt nutzen) werden diese Bruchstücke zu ganzen Wort-Segmenten zusammengesetzt⁶⁷ und redundant im Transkript notiert (einmal in Bruchstücken, einmal am Stück, ferner als Bestandteil ganzer Redebeiträge usw.). Nur wenn die Wörter unzerlegt vorliegen, lassen sich die Transkripte effizient durchsuchen. Für das Retrieval wären darüber hinaus die Wortstämme hilfreich, denn in den Transkripten gemäß der *segmented-transcription.dtd* werden als Wörter nur die *tokens*, d.h. Wörter, wie sie im Transkripttext geschrieben wurden, flektiert oder auch dialektal abweichend von der Standardorthographie notiert. Für die Suche wären aber deren *types* (Grundwörter) hilfreicher. In EXMARaLDA könnte man Grundwörter oder Wortstämme manuell⁶⁸ in eigenen Wortstamm-Zeilen (mit einem eigenen Typ) parallel unter den Sprecher-Zeilen mit den flektierten oder unvollständigen Wortformen des Transkripts notieren. Die Nutzung dieser speziellen Zeilen wäre dann ebenfalls "manuell" zu organisieren (ausblenden in der Präsentation, einblenden für Volltext-Retrieval).

Bei zusammengesetzten Wörtern, deren Bestandteile im propositionalen Aufbau nicht zusammenhängend realisiert werden - eine für das Deutsche typische Klammerstruktur - (z.B. "vorsehen", und: "ich *sehe* das für dich *vor*"), gibt es bei der Kennzeichnung noch ein besonderes Problem. "Fortsetzungsereignisse" gibt es in EXMARaLDA nicht. Die Zusammengehörigkeit mehrerer Teile wäre technisch nur dadurch zu dokumentieren, dass in solchen Fällen Kennungen explizit

⁶⁶ Das Layout solcher Kombinationen wird typographisch oft nicht zufriedenstellend realisiert. Das ist aber den Routinen des Betriebssystems anzulasten, mit denen die Positionierung der Diakritika über den Grundbuchstaben erfolgt.

⁶⁷ Es ist angesichts der zahlreichen Sonderzeichen nicht ganz trivial, Wortgrenzen zu identifizieren. Man wird künftig die Zeichen, die in EXMARaLDA als Worttrenner fungieren, und die Symbole zur Kennzeichnung des Endes von Äußerungen vorgeben können.

⁶⁸ Automaten dürften bei der Lemmatisierung von Transkriptionen gesprochener Sprache nicht sonderlich zuverlässig sein.

manuell vergeben werden (in weiteren Zeilen), da die internen Kennungen für die Zellen automatisch erzeugt werden und sich bei Änderungen des Transkripts automatisch ändern können. Referenzen auf automatische interne Kennungen würden dabei aber unberücksichtigt bleiben. Für grammatikalische Analysen sind solche Kennzeichnungen mehrgliedriger Textsequenzen unverzichtbar. EXMARaLDA bietet hierfür keine effiziente Unterstützung.

2.6. Positionierung und Navigation in EXMARaLDA-Transkripten

EXMARaLDA kennt keine äußerlichen Annotationen zur groben Gliederung eines Transkripts. Für Gliederungszwecke und zur Positionierung im Transkript eignen sich die Zellengrenzen. Sie sind strukturell auf der Zeitachse verankert und funktionieren ähnlich wie die Synchronisationspunkte im syncWRITER-Programm. Die Zeitanker werden automatisch fortlaufend durchgezählt. Wenn im Zuge der Verfeinerung des Transkripts weitere Zeitanker eingefügt werden, ändern sich die Nummern der jeweils nachfolgenden. Man kann jedoch jeden Zeitanker auch mit einer Zeitangabe versehen. Solche Angaben bleiben bei anderen Änderungen des Transkripts erhalten.

In den exportierten Präsentationsformen werden auch die Partiturblöcke durchgezählt. Diese Zählung entsteht erst beim Umbruch (in Abhängigkeit von der Zeilenlänge) während des Exports und ist beim Transkribieren im Editor nicht zu sehen. Zeilenzählungen der Partiturzeilen (wie für IDS-DIDA) werden beim Export nicht erzeugt.

2.7. Integration von Medien

Die logische Struktur von Folgen von Zeitpunkten entspricht dem Grundkonzept der Annotations-Graphen und könnte deshalb im EXMARaLDA-Konzept in naheliegender Weise eindeutig auf technische Strukturen abgebildet werden. Verwenden lassen sich die optimal zum EXMARaLDA-Konzept passenden zeitabhängigen Strukturen im Editor derzeit aber nicht so recht. Man kann nämlich die Ton- oder Video-Medien mit EXMARaLDA nicht effizient beim Transkribieren nutzen, weder um Synchronanker zu setzen (d.h. Zellen zu zerlegen), noch um damit synchron in Aufnahme und Transkript zu navigieren. Die Stoppuhr, das Metronom oder externe Programme bleiben leider unverzichtbar. Die Verknüpfung der Transkripte mit den digitalisierten Tonaufnahmen oder Videos ist unvollständig realisiert und das Konzept ist noch nicht vollständig erkennbar. In den Datenstrukturen sind Partiturzeilen für Verknüpfungen mit anderen Dateien vorgesehen, z.B. mit Tondateien. Diese Verknüpfungen kann man im Editor jetzt schon herstellen. Dabei werden Zellen definiert, deren Reichweiten zu den Längen der verknüpften Mediendateien passen müssen. Im einfachsten Fall füllt die Zelle mit der Verknüpfung die komplette Transkriptionszeile aus, reicht also vom Anfang bis zum Ende eines Transkripts. Es können aber auch mehrere Ausschnitte einer Aufnahme in mehreren Zellen mit dem Transkript verknüpft werden. Ein Verweis von einer Zelle im Transkript auf eine externe Mediendatei stellt zunächst nur eine recht grobe Verknüpfung her, nämlich von einer Stelle im Transkript zum Anfang einer Aufnahme. Nur wenn zu allen relevanten Zellen des

Transkripts auch die Zeitpunkte (in Sekunden) für den Zellenanfang und das Zellenende notiert wurden, ließen sich die Aufnahmen überhaupt so differenziert nutzen, wie das Transkript es vorsieht, d.h. wie es auf Zellen verteilt vorliegt.⁶⁹ Wie die Wiedergabe ablaufen soll, wenn mehrere Verknüpfungen mit Mediendateien definiert wurden, deren Reichweiten sich überlappen, (das kann bei mehrspurigen Aufnahmen aus verschiedenen Kamerapositionen durchaus sinnvoll sein), bleibt konzeptionell unklar. Eigentlich müssten dann mehrere Wiedergabeprogramme parallel betrieben werden. Die Kontrolle der Konsistenz der Verknüpfungen mit Mediendateien (Lücken oder Überdeckungen?) bleibt in jedem Falle dem Anwender überlassen.

Da es kein integriertes Wiedergabeprogramm gibt, ist während des Transkribierens von digitalisierten Medien also weiterhin der lästige Wechsel zwischen dem Transkriptionseditor und einem getrennten Wiedergabeprogramm⁷⁰ erforderlich. Immerhin können in den EXMARaLDA-Datenstrukturen alle erforderlichen Informationen untergebracht werden, daher sollte der Zugriff auf digitalisierte Medien in den Editor integrierbar sein, ohne das Konzept der Datenstrukturen abwandeln zu müssen.⁷¹ Erst nach einem Export der Transkripte z.B. ins HTML-Format lassen sich die Verknüpfungen zu Mediendateien praktisch verwenden, wenn auch nur so grob wie oben beschrieben (die Wiedergabe beginnt immer am Anfang der Datei, auf die verwiesen wird). Auch nach dem Export in das TASX-System oder in praat lassen sich die Tondateien zusammen mit dem Transkript nutzen (es ist hier sogar ein gezielterer Zugriff auf einzelne Stellen in den Tondateien möglich), das Layout der Transkripte weicht in diesen Systemen jedoch stark von der ursprünglichen Form ab.

2.8. Übernahme vorhandener Transkripte in EXMARaLDA

Der *Import* von Transkripten, die mit anderen Transkriptionseditoren erstellt wurden, ist derzeit über folgende Import-Filter möglich:

- Für das strukturarme Simple-EXMARaLDA-Format im Listenformat. Es besteht aus schlichten Textzeilen, die mit der Sprechersigle eingeleitet werden.

⁶⁹ Wie eng ein Alignment von Transkript und Tonaufnahme auch bei umfangreichen Transkripten heute sein kann, zeigt das DSAv im IDS anhand einiger über das Internet öffentlich zugänglicher Transkripte. Diese Transkripte wurden wortweise aligniert. Siehe unter www.IDS-Mannheim.de/DSAv.

⁷⁰ Wie die Integration aus Benutzersicht aussehen könnte, demonstriert mein Ton-und-Text-Programm. Es läßt sich "solo" neben einem beliebigen Transkriptionseditor unter Windows nutzen, vor allem aber "synchron" zum Transkribieren in Word integrieren. Das ist technisch gesehen eher ein Kompromiss und keine vorbildliche Lösung, aber es lassen sich damit immerhin auf sehr bequeme Weise Positionen in den Aufnahmen an den Transkripttext binden und zum Text gehörige Ausschnitte aus Tonaufnahmen aus Word heraus wiedergeben. Siehe unter: www.Ton-und-Text.de.

⁷¹ Eventuell kann man einige Anleihen beim TASX-Annotator machen, der dafür entwickelt wurde, Mediendateien mit Transkripttext zu annotieren, also für Verknüpfungen in der anderen Richtung. Es ist "im Prinzip" ein Im- und Export für TASX vorbereitet. Der Datenaustausch aus EXMARaLDA in den Annotator (Version 1.0, alpha, vom 29.11.2001 mit Update vom 20.6.2002) gelang nur mangelhaft. Die Datenstrukturen aus EXMARaLDA wurden in den TASX-Strukturen nicht adäquat abgebildet. Das Problem könnte aber bald behoben sein. Siehe unter: www.tasxforce.Uni-Bielefeld.de.

Aus jeder Zeile wird im EXMARaLDA-System eine Zelle. Es können Synchronanker gekennzeichnet werden, jedoch ohne Zeitangaben, und es lassen sich verbale und nonverbale Äußerungen oder Kommentare unterscheiden. Das Format eignet sich recht gut, Transkriptionen, die nach den GAT-Konventionen erstellt wurden, zu importieren. Die "Freitext"-Transkripte nach GAT müssen vorher konsequent in das Simple-EXMARaLDA-Format gebracht werden. Das erfordert manuelle Vorbereitungen, insbesondere an den Stellen gleichzeitigen Sprechens und in vielen Fällen auch bei der Verwendung von Sonderzeichen.⁷²

- TASX-Transkripte sind strukturell ähnlich wie EXMARaLDA-Strukturen aufgebaut, sie enthalten insbesondere Zeitangaben für den Anfang und das Ende jedes "events". Das läßt sich gut auf Zellen in EXMARaLDA abbilden.
- Transkripte, die in TextGrids des praat-Systems erzeugt wurden, sind ebenfalls so aufgebaut, dass sie sich gut in EXMARaLDA-Strukturen überführen lassen (TextGrids sind "Zellen in Zeilen").
- Für Transkripte, die mit HIAT-DOS erstellt wurden, gibt es einen "Import"-Filter, der seinen Namen nur mit einschränkenden Anführungsstrichen trägt. Der Filter versucht HIAT-DOS-Transkripte in Zellen zu zerlegen. Das lässt sich aber prinzipiell mit keinem Automaten völlig zufriedenstellend erledigen. Transkripttext, Intonation und NVK-Annotationen werden in HIAT-DOS mit fixlaufenden Leerzeichen passend zueinander positioniert. Passagen, in denen sich die Sprecher "diszipliniert" abwechseln und die nicht weiter annotiert wurden, sind unproblematisch. Bei Passagen gleichzeitigen Sprechens ist es aber nicht immer zweifelsfrei möglich, automatisch zu einer brauchbaren Zerlegung in Zellen zu gelangen. Nach dem Import ist genaue Kontrolle und eventuell manuelle Nacharbeit erforderlich.⁷³

EXMARaLDA wurde am SFB 538 als Nachfolger für das syncWRITER-Programm konzipiert. Da das binäre syncWRITER-Format von den früheren Entwicklern bisher nicht offengelegt wurde (und sich nur mühsam entschlüsseln ließe), gibt es für solche Transkripte derzeit keinen funktionierenden Import. Die Strukturen der beiden Partitureditoren wurden nach ähnlichen Konzepten entwickelt (sie werden jedoch völlig unterschiedlich kodiert). Deshalb werden sich, wenn denn die binäre Form offengelegt oder entschlüsselt ist, syncWRITER-Transkriptionen auch fast verlustfrei in EXMARaLDA-Transkriptionen wandeln

⁷² Die Verwendung von Sonderzeichen in GAT-Transkripten ist nur bis auf die Ebene der Glyphen geregelt, nicht aber auf der Ebene der Kodierung. Hochgestellte Pfeile z.B. müssen letztlich in geeignete Unicode-Zeichen umgesetzt werden. Das alles wird umständlich bleiben, weil es Transkript-spezifische temporäre Ersatzdarstellungen notwendig macht (zu deren Handhabung weitere Programme oder kleine Skripte erforderlich sind) bis die Transkripte in EXMARaLDA in der gewünschten Form kodiert sind.

⁷³ Leider wurde der "Import"-Filter (in EXMARaLDA-Version 1.2.2) nur für die am SFB 538 vorhandenen HIAT-DOS-Transkripte entwickelt. Wenn jedoch Intonationszeilen vorliegen, wird die Partiturstruktur zerstört. Unterstreichungen fehlen nicht nur (zu den prinzipiellen Problemen Annotationen des Strukturtyps A in EXMARaLDA zu verarbeiten siehe oben unter 2.3.1), die unterstrichenen Teile, die in HIAT-DOS (intern) kodiert sind, werden unkonvertiert übernommen und sind nun nicht mehr lesbar. Außerdem gehen einige Abstände zwischen dem letzten Wort einer HIAT-DOS-Zeile und dem ersten Wort am Anfang der folgenden verloren.

lassen.⁷⁴ IDS-DIDA-Transkripte sind strukturell ähnlich wie HIAT-DOS. Beide verwenden fixlaufende Schriften und stützen die Synchronität mehrerer Sprecherzeilen auf die identische Anzahl der Zeichen in allen Zeilen, gemessen vom Anfang des Transkripts. Abgrenzungen von Redebeiträgen erfolgen implizit durch Leerstellen in den Sprechpausen. Explizite Abgrenzungen dienen nach den DIDA-Konventionen allenfalls dazu, gleichzeitig Gesprochenes etwas präziser zu kennzeichnen. Ein Importfilter für DIDA-Transkripte hätte also ähnliche Schwierigkeiten wie der für HIAT-DOS-Transkripte zu bewältigen.

Für alle oben beschriebenen Transkript-Importe gilt, dass sie den vorliegenden Transkripttext mehr oder weniger effizient und treffsicher in EXMARaLDAs Zeilen kopieren. Annotierende Textattribute und intralineare Annotationen (Strukturtypen A und B, siehe 2.3.1 und 2.3.2) werden dadurch nicht explizit gemacht, also auch nicht kodiert. Man wird später mit XML-Werkzeugen nicht effizienter darauf zugreifen können als auf Annotationen in unformatierten Textdateien. Um die Annotationen in EXMARaLDA explizit zu kodieren, wäre es erforderlich, auch die gewohnte Präsentationsform der Transkripte deutlich abzuwandeln. Dann entsprächen die Transkripte aber nicht mehr den Konventionen, die für sie beim Transkribieren noch galten, oder man müßte die Annotationen redundant notieren, einmal wie bisher zur Präsentation und einmal (unsichtbar) zur Kodierung in XML-Strukturen. Die redundante Notation wäre aufwändig und fehlerträchtig, also keine brauchbare generelle Empfehlung.

2.9. Anpassung von EXMARaLDA an die Konventionen oder umgekehrt?

Ein Teil der Transkribenten wird an den von ihnen langjährig verwendeten, bewährten Konventionen - und den nur bedingt tauglichen Werkzeugen, sie zu realisieren - festhalten und das EXMARaLDA-Programm vorerst nicht verwenden. Sie werden auf einen Fortschritt bei der Entwicklung von Transkriptionseditoren hoffen - ob der EXMARaLDA-2 oder völlig anders heißt, ist egal -, weil sie einen Editor brauchen, der mehr Strukturen für das Annotieren vorsieht und der zu der von ihnen bevorzugten Konvention exakt passt.

Auch wenn EXMARaLDA nicht in der Lage ist, alle Details etablierter Transkriptionskonventionen konform zu realisieren, wird ein anderer Teil der Transkribenten den technischen Komfort, den EXMARaLDA bietet, nutzen und dabei die Inkompatibilitäten zu den Konventionen ignorieren. Sie werden mit Ersatzdarstellungen und Kompromisslösungen versuchen, ihre Annotate "unterzubringen" und dabei für identische Phänomene (siehe z.B. die Alternativen unter 2.3.2) zu individuellen, wiederum inkompatiblen Lösungen kommen. Um bei solchen Fragen keinen Wildwuchs mit neuen Inkompatibilitäten aufkommen zu lassen, wäre es erforderlich, detaillierte Anleitungen für die "gültige" Handhabung von EXMARaLDA zu verwenden, die den Umgang mit dem Werkzeug konform zu den Konventionen erläutern. Die EXMARaLDA-Dokumentation kann das

⁷⁴ Probleme mit einigen speziellen MAC-Zeichensätzen (z.B. HIAT-Times) dürften zu bewältigen sein.

nicht bieten, weil es sich um spezifische Fragen der Konventionen handelt.⁷⁵ Es wären für eine solche Anleitung zur Nutzung von EXMARaLDA wohl auch kleinere Anpassungen der Konventionen in den Fällen erforderlich, die sich mit EXMARaLDA nicht wie gewohnt realisieren lassen.

2.10. Fazit zu EXMARaLDA

EXMARaLDA ist ein System, das zahlreiche neuere Konzepte und Techniken verwendet. Gegenüber den bisher verfügbaren Transkriptionswerkzeugen bietet es vor allem die Offenlegung und - damit verbunden - systematischere Datenstrukturen. Es ist jedoch vor allem wegen der eingeschränkten Annotationsmöglichkeiten konzeptionell nicht flexibel genug, um die älteren DV-technisch mangelhaften, aber rezeptions-ökonomisch elaborierten Transkriptformate verlustfrei abzulösen.⁷⁶ Der Transkriptionseditor kann konzeptionell natürlich nicht flexibler sein als es seine grundlegenden Datenstrukturen vorsehen, aber er schöpft derzeit die Möglichkeiten, die in den DTDs stecken, noch bei weitem nicht aus.

EXMARaLDA konzentriert sich sehr auf den Bedarf am SFB 538, auch wenn es als allgemein verwendbares Transkriptionswerkzeug angeboten (und erwartet) wird. Im Falle von Kompromissen, die wohl aufgrund der begrenzten Entwicklungskapazitäten notwendig waren, wurde verständlicherweise pro domo entschieden. Das sieht man dem Endergebnis in den Details leider an, und es würde sich bei anderen Nutzern auswirken: teils als SFB 538-Besonderheit (z.B. wird mit jedem Sprecher "L2" notiert) und teils als lästige Umständlichkeit oder Unmöglichkeit, beim Transkribieren nach gewohnten und bewährten Konventionen zu transkribieren (Hervorhebungen lassen sich nur recht umständlich einbauen, Unterstreichungen fehlen völlig; die Notation der Intonation bleibt ein kryptischer Kompromiss wie bisher schon mit syncWRITER, usw.). Die Datenformate von EXMARaLDA sind nicht so flexibel, dass sie die verbreiteten Transkriptionssysteme DV-technisch abbilden könnten und sie konsequent auf eine höhere DV-technische Stufe brächten - obwohl EXMARaLDAs Grundkonzept dazu einige moderne Voraussetzungen mitbringt. Es ist nur teilweise möglich, die bekannten Transkriptionssysteme mit EXMARaLDA nachzustellen und zugleich die Annotate in XML-Strukturen zu notieren.

Auch die differenziertesten Strukturen, die für Segmentierungen im Prinzip im EXMARaLDA-System verfügbar wären, kann der EXMARaLDA-Editor, der für Anwender die einzig praktisch verwendbare Schnittstelle zu den Transkriptionsdaten darstellt, nicht erfassen und wie gewohnt anzeigen oder exportieren. Der Editor operiert nur auf einer spezialisierten Teilmenge der EXMARaLDA-Daten. Eine *konkrete technische* Umsetzung des Konzepts der Annotationsgraphen (das sind wie gesagt *abstrakte logische* Strukturen) erfordert im Detail noch einige

⁷⁵ Für syncWRITER gab es ein Bedienungshandbuch des Herstellers und von den "Erfindern" des syncWRITER und außerdem ein "Manual für das Transkribieren mit dem Programm syncWRITER nach dem Verfahren der Halbinterpretativen Arbeitstranskriptionen (HIAT)" (Rehbein et. al 1993).

⁷⁶ Der Begriff Transkriptformat ist hier doppeldeutig: er bezeichnet das (interne) Datenformat für Transkripte und das (äußerliche) Präsentationsformat. Das ist der wesentliche Mangel der bisher verwendeten Werkzeuge, die aus den äußeren Darstellungsformen der Konventionen, die sie umsetzen sollen, umweglos die inneren Formate von Transkripten herleiten.

andere Datenstrukturen und diverse weitere Erfassungs- und Darstellungsmittel, um damit ebenso effizient Transkripte erstellen und rezipieren zu können wie bisher mit den etablierten Werkzeugen, mit denen zu Recht kaum jemand rundum zufrieden ist.

Der EXMARaLDA-Editor funktioniert als unkomplizierter plattformunabhängiger Partitur-Editor,⁷⁷ wird aber seinem Namen leider nur teilweise gerecht, weil er die *discourse annotation* (DA) nur unvollständig (zumindest nicht in der gewohnten Präsentationsform) mittels *markup* (MAR) ermöglicht. Transkripte damit zu erstellen garantiert deren Austauschbarkeit in etwa so zuverlässig - sogar längerfristiger -, wie den Austausch von Transkriptionen, die mit Word erstellt wurden. Die technische Grundlage ist in beiden Fällen vom Hersteller geregelt und weit weg von den Annotaten in den Transkripten. EXMARaLDA "merkt" von den meisten Annotationen ebensowenig wie Word und kann unter dieser Bedingung nicht mehr sein als ein Editor für das Schreiben in parallelen segmentierbaren Zeilen. Einen Editor für Transkriptionen in einem XML-Format zu verwenden, aber ohne das auch mit dem Namen versprochene MARKup an den eigentlich spannenden Stellen einzusetzen, wäre paradox. Mit EXMARaLDA muss man jedoch in einigen Details anders transkribieren als bisher durch die Transkriptionskonventionen vorgesehen, wenn man das EXMARaLDA-Konzept konsequent nutzen möchte. Annotationen mit dem Editor jedoch nur präsentationsfähig zu notieren, sie aber nicht in den Strukturen festzuhalten, würde das innovative EXMARaLDA-Konzept unterlaufen. Ein bisschen fortschrittlich werden und XML-Strukturen verwenden, aber ansonsten in den Annotationsstrukturen der Transkripte alles so zu lassen, wie es immer schon war, wäre inkonsequent.

3. Nachbemerkung

Annotationen sind Metatexte in Sprachdaten, die Resultate von Analysen sind. Die wissenschaftliche Arbeit, die in ihnen steckt, soll vielfältig - auch von anderen - genutzt und langfristig konserviert werden. Das aber funktioniert mit vertretbarem Aufwand praktisch nur dann, wenn man sich vorher auf recht formale Entscheidungen einläßt, wie mit der Erstellung einer DTD (oder vergleichbarer Definitionen) verbunden sind. Die Anforderungen, die von Transkribenten an Transkripte gestellt wurden, blieben bisher jedoch meist (im Wortsinne) oberflächlich, sie bezogen sich auf die sichtbaren Erscheinungsformen von Transkriptionen. Die Konzeption adäquater technischer Grundlagen für sprachwissenschaftliche Transkriptionen blieben in Deutschland bisher überwiegend Programmierern, bestenfalls Informatikern⁷⁸ überlassen. Überlegungen zu Datenstrukturen, wie sie hier angestellt wurden, müssen aber von Gesprächsforschern selber vorgenommen werden, weil mit diesen Strukturen über die technische Qualität ihrer Transkripte entschieden wird. Und die wirkt sich ganz praktisch aus auf den Umgang mit ihren Erzeugnissen: Ob sich Analysen in Transkripten zutreffend in Annotaten

⁷⁷ Ich selber habe den Editor nur mit Java unter Windows getestet.

⁷⁸ Es geht bei dieser Unterscheidung nicht um akademische Titel, sondern um die Art der Herangehensweise der Entwickler: handwerklich vs. wissenschaftlich. EXMARaLDAs Konzept wurde wissenschaftlich begründet (siehe Schmidt (2000a)).

festhalten, präzise adressieren und differenziert recherchieren lassen, oder ob sich Transkripte auch mit anderen Werkzeugen als mit denen, die bei ihrer Erstellung eingesetzt wurden, verwenden lassen und wie effizient all dies zu erledigen ist, hängt wesentlich ab von den zugrundeliegenden Datenstrukturen.

Unter den auch nur näherungsweise einschlägigen Publikationen findet man nur sehr wenige deutschsprachige.⁷⁹ Diese Zurückhaltung hierzulande ist erstaunlich angesichts der langen Tradition, der elaborierten Transkriptionskonventionen und der Bedeutung der Fragen nach den für Transkripte adäquaten Datenstrukturen. In anderen Ländern gibt es seit längerer Zeit Bemühungen von Sprachwissenschaftlern, sich an diesen Überlegungen nicht nur zu beteiligen, sondern sie aktiv voran zu bringen.⁸⁰ Die daraus hervorgegangenen Datenformate sind nicht unbedingt überzeugend - auch die jüngeren nicht -, aber man sieht immerhin, dass dort an diesen Fragen intensiv gearbeitet wird.

Solange sich Gesprächsforscher nicht ebenso engagiert und detailliert an den DV-technischen Spezifikationen beteiligen, wie sie dies bei der Definition von Transkriptionskonventionen taten bzw. immer noch tun, wird es auch kein geeignetes Datenformat für Transkriptionen geben, das gleichzeitig in möglichst vielen unterschiedlichen Projekten mit unterschiedlichen Annotationsbedürfnissen einsetzbar ist. Und es wird - darauf aufbauend - solange auch kein Transkriptionswerkzeug geben, das geeignet wäre, damit nach ähnlichen, im Detail aber unterschiedlichen Konventionen zu transkribieren und Transkripte untereinander auszutauschen - vorausgesetzt, die Unterschiede in den Konventionen beziehen sich nur auf die präsentierte Oberfläche der Transkripte.⁸¹ Nur ein großer Interessentenkreis, der das Werkzeug in unterschiedlichen Projekten verwenden könnte, rechtfertigt den Einsatz der nicht unerheblichen Mittel, ein solch flexibles Werkzeug nach professionellen Maßstäben zu erstellen, kontinuierlich zu pflegen und weiterzuentwickeln. EXMARaLDA geht einen wichtigen Schritt in diese Richtung und zeigt, wie man ein solches Vorhaben im Prinzip angehen könnte, tut dies aber nicht konsequent genug.

⁷⁹ Z.B. im Buch von Lobin (Hg.) 1999. Einige Publikationen stammen aus der DFG-Forscherguppe 'Texttechnologische Informationsmodellierung', in der Konzepte zur Strukturierung linguistischer Daten entwickelt werden. Innerhalb der Gruppe befasst sich insbesondere das Projekt 'Sekimo' (Sekundäre Informationsstrukturierung und vergleichende Diskursanalyse) auch mit gesprochener Sprache, siehe unter www.text-technology.de.

⁸⁰ Einige exemplarische Projekte: AG am LDC (Linguistic Data Consortium, USA), LACITO am CNRS (Centre National de la Recherche Scientifique, Frankreich) oder Talkbank am LDC und CMU (Carnegie Mellon University, USA). Einige dieser Bemühungen zur Entwicklung von Techniken für den Aufbau von Korpora gesprochener Sprache werden in der Spoken Language Group des ISLE-Projekts (International Standards for Language Engineering) koordiniert.

⁸¹ Es kam sogar schon das Missverständnis auf, dass EXMARaLDA als Konverter für die bekannten Transkriptionssysteme fungieren könnte ("Transkribierter Text kann vielmehr auf der Grundlage von XML [...] so gestaltet werden, dass er letztlich zunächst unabhängig vom Transkriptionssystem bzw. -format zusammengestellt bzw. erstellt und dann in jedem beliebigen Transkriptionssystem (HIAT, GAT, CHAT usw.) dargestellt [...] werden kann", Dittmar 2002:218). Dieser Wunsch wird mit EXMARaLDA nicht in Erfüllung gehen - und seine Entwickler hatten wohl auch nicht vor, ein solches Transkriptions-Chamäleon zu erzeugen.

4. Literatur

- Bird, Steven / Liberman, Mark (2001): A formal framework for linguistic annotation. In: *Speech Communication* 33 (1,2), European Association for Signal Processing, North-Holland Publ. Comp., Amsterdam, 23-60.
- Dittmar, Norbert (2002): *Transkription*, 1. Auflage, Opladen: Leske + Budrich.
- Ehlich, Konrad / Rehbein, Jochen (1976): Halbinterpretative Arbeitstranskriptionen (HIAT). In: *Linguistische Berichte* 45, 21-41.
- Ehlich, Konrad (1979): Erweiterte halbinterpretative Arbeitstranskriptionen (HIAT2); Intonation. In: *Linguistische Berichte*, 59, 51-75.
- Ehlich, Konrad (1981a): Zur Notierung nonverbaler Kommunikation für diskursanalytische Zwecke (Erweiterte halbinterpretative Arbeitstranskriptionen HIAT 2). In: Winkler, P. (Hg.) *Methoden der Analyse von Face-to-Face-Situationen*, Stuttgart: Metzler, 302-329.
- Ehlich, Konrad (1981b): Die Wiedergabe intonatorischer, nonverbaler und aktionaler Phänomene im Verfahren HIAT. In: Lange-Seidl, A. (Hg.) *Zeichenkonstitution*, Bd. 2 (Akten des 2. Semiotischen Kolloquiums), Berlin: de Gruyter, 174-186.
- Ehlich, Konrad (1992a): Computergestütztes Transkribieren - das Verfahren HIAT-DOS. In: Richter, Günther (Hg.) (1992): *Methodische Grundfragen der Erforschung gesprochener Sprache*, Frankfurt a. M.: P. Lang, 47-59.
- Ehlich, Konrad (1992b) HIAT - a Transcription System for Discourse Data. In: Edwards, Jane A.; Lampert, Martin D. (eds.) (1992): *Talking Data: Transcription and Coding in Discourse Research*, Hillsdale, NJ.: L. Erlbaum Ass., 123-148.
- Genette, Gérard (2001): *Paratexte: Das Buch vom Beiwerk des Buches*. Frankfurt a. M.: Suhrkamp.
- Glas, Reinhold / Ehlich, Konrad (2000): Deutsche Transkripte 1950 bis 1995. Ein Repertorium. In: *Arbeiten zur Mehrsprachigkeit* 63 / 2000, Institut für Germanistik I / Arbeitsbereich Deutsch als Fremdsprache, Arbeitsstelle Mehrsprachigkeit / Research Center for Multilingualism, Universität Hamburg, Hamburg, 2000. (Die HTML-Version ist unter www.DaF.Uni-Muenchen.de/DTR zu finden.)
- Lenk, Uta (1999): Notation systems in spoken language corpora. In: Verschueren, Jef; Östmann, Jan-Ola; Blommaert, Jan; Bulcaen, Chris (eds.) (1999): *Handbook of Pragmatics 1999*, Amsterdam, Philadelphia: John Benjamins Publ. Co.
- Lobin, Henning (Hg.) (1999): *Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und Hypertext Engineering*. Wiesbaden: Westdeutscher Verlag.
- Redder, Angelika (2001): Aufbau und Gestaltung von Transkriptionssystemen. Artikel 100. In: Brinker, Klaus; Antos, Gerd; Heinemann, Wolfgang; Sager Sven F. (Hrsg.): *Text- und Gesprächslinguistik (HSK 16.2)*. Berlin, New York: Walter de Gruyter, 1038-1059.
- Redder, Angelika (2002): Professionelles Transkribieren. In: Jäger, Ludwig; Stanzek, Georg (Hrsg.): *Transkribieren*. München: Fink, 115-131.
- Rehbein, Jochen / Griebhaber, Wilhelm / Löning, Petra / Hartung, Marion / Bührig, Kristin (1993): *Manual für das computergestützte Transkribieren mit*

- dem Programm SyncWRITER nach dem Verfahren der Halbinterpretativen Arbeitstranskriptionen (HIAT). Hamburg.
- Rehbein, Jochen (1995): Segmentieren, Memo 64 des Verbundprojektes Verbomobil. Hamburg
- Schmidt, Thomas (2002a): EXMARaLDA - ein System zur Diskurstranskription auf dem Computer. In: Arbeiten zur Mehrsprachigkeit, Serie B (34). Hamburg.
- Schmidt, Thomas (2002b): Gesprächstranskription auf dem Computer - das System EXMARaLDA. In: Gesprächsforschung - Onlinezeitschrift zur verbalen Interaktion 3, 1-23 (www.gespraechsforschung-online.de).
- Schneider (2002): Kritische Stellungnahme zur Darstellung DV-technischer Aspekte in Norbert Dittmars Buch "Transkription". In: Gesprächsforschung - Onlinezeitschrift zur verbalen Interaktion 3, 38-43 (www.gespraechsforschung-online.de).
- Selting, Margret et al. (1998): Gesprächsanalytisches Transkriptionssystem (GAT). In: Linguistische Berichte 173, 91-122.
- Sperberg-McQueen, C. M. / Burnard, Lou (eds.) (1994): Guidelines for Electronic Text Encoding and Interchange (TEI P3). Chicago, Oxford: Text Encoding Initiative.

5. URLs

- AG (Annotation Graphs, am LDC)
agtk.sourceforge.net
- AG am LDC (Linguistic Data Consortium, USA),
www ldc .edu/bird
- BAS (Bayerisches Archiv für Sprachsignale)
www.phonetik.Uni-Muenchen.de/Bas/BasHomedeu.html
- CHAT (Codes for the Human Analysis of Transcripts) und
CLAN (Computerized Language ANalysis)
childes.psy.cmu.edu
- DTR (Deutsche Transkripte 1950 bis 1995. Ein Repertorium)
www.DaF.Uni-Muenchen.de/DTR
- EXMARaLDA (EXtensible MARKup Language for Discourse Annotation)
www.rrz.Uni-Hamburg.de/exmaralda
- GAT (Gesprächsanalytisches Transkriptionssystem)
www.fbfs.Uni-Hannover.de/sdls/schlobi/schrift/GAT
- HIAT und HIAT-DOS (Halbinterpretative Arbeitstranskription)
www.DaF.Uni-Muenchen.de/HIAT
- IDS-DIDA (Institut für Deutsche Sprache, Diskurs-Datenbank)
www.IDS-Mannheim.de/prag/dida
- IDS-DSAv (Institut für Deutsche Sprache, Deutsches Spracharchiv)
www.IDS-Mannheim.de/DSAv
- ISLE (International Standard for Language Engineering)
www.mpi.nl/ISLE
- LACITO (Langues et Civilisationsà Tradition Orale, am Centre National de la
Recherche Scientifique, Frankreich)
lacito.vjf.cnrs.fr
- LDC (Linguistic Data Consortium at the University of Pennsylvania)
www ldc .upenn.edu

OLAC (Open Language Archive Consortium)

www.OLAC.org

TalkBank (am LDC und an der Carnegie Mellon University)

www.talkbank.org

TASX (Time Aligned Signal data eXchange)

tasxforce.Uni-Bielefeld.de

TEI (Text Encoding Initiative)

www.tei-c.org

Texttechnologische Informationsmodellierung (DFG-Forschergruppe)

www.text-technology.de

TuT (Ton und Text)

www.Ton-und-Text.de

6. Kommentiertes Demo-Transkript im EXMARaLDA-Format

Nachfolgend wird das Demo-Transkript zu 2.1.3 im EXMARaLDA-eigenen Datenformat der *basic-transcription.dtd* wiedergegeben. Das vollständige XML-formatige Transkript dient zum einen als Beleg für einige Aussagen im Text, zum anderen illustriert es, wie im EXMARaLDA-System Annotationsstrukturen mit XML-Konstrukten festgehalten werden. Das Transkript wird um Lesehilfen ergänzt, um XML-Laien die Orientierung zu ermöglichen.

Einleitung einer XML-Datei mit Angabe des verwendeten Zeichensatzes "UTF-8":

```
<?xml version="1.0" encoding="UTF-8"?>
```

Copyright-Notiz des EXMARaLDA-Entwicklers als XML-Kommentar:

```
<!-- (c) 2001 by Thomas Schmidt (thomas.schmidt@uni-hamburg.de) -->
```

Die nachfolgenden abgrenzenden Einträge werden paarig notiert. An erster übergreifender Position steht der Name der zuständigen DTD.

```
<basic-transcription>
```

Hier beginnt der Transkriptionskopf:

```
<head>
```

Darin gibt es diverse Meta-Informationen:

```
<meta-information>
```

Eine typische XML-Kodierung: die eigentliche Information, hier "Demo-Transkripte", steht zwischen zwei Klammern, hier "project-name":

```
<project-name>Demo-Transkripte</project-name>
```

```
<transcription-name>ANNODEMO</transcription-name>
```

```
<referenced-file url="ANNODEMO.WAV"/>
```

Ab hier der Block der Benutzer-definierten Informationen des Kopfes:

```
<ud-meta-information>
```

```
<ud-information attribute-name="Copyright">Schneider (Dortmund)</ud-information>
```

```
<ud-information attribute-name="Aufnahmedatum">2002-05-01</ud-information>
```

```
<ud-information attribute-name="Aufnahmeort">Dortmund</ud-information>
```

```
<ud-information attribute-name="Transkribent">Schneider</ud-information>
```

```
<ud-information attribute-name="Transkriptdatum">2002-12-20</ud-information>
```

```
</ud-meta-information>
```

Kommentar in den Meta-Informationen (nicht zu verwechseln mit einem XML-Kommentar, s.o.):

```
<comment>Transkript zur Illustration von Annotationsstrukturen.</comment>
```

```
<transcription-convention>HIAT</transcription-convention>
```

```
</meta-information>
```

Ab hier die Sprechertabelle:

```
<speakertable>
```

Jeder Sprecher erhält ein (internes) "id"-Attribut (dieser hier mit dem Inhalt "SPK0"):

```
<speaker id="SPK1">
```

Jeder Sprecher erhält eine Sigle:

```
<abbreviation>S1</abbreviation>
```

... und hat ein Geschlecht:

```
<sex value="m"/>
```

... und spricht eine Sprache:

```
<languages-used><language lang="de"/></languages-used>
```

L1 und L2 müssen rein, auch wenn sie leer sind:

```
<l1><language lang="de"/></l1>
```

```
<l2></l2>
```

Auch Sprecher-Informationen können vom Benutzer definiert werden:

```
<ud-speaker-information>
```

Der Inhalt des Attributs "attribute-name" (hier "Alter") ist hier der eigentliche Name der Information, der Inhalt zwischen den Klammern "ud-information" die Information ("28"):

```
<ud-information attribute-name="Alter">28</ud-information>
```

```
<ud-information attribute-name="Geburtsort">Essen</ud-information>
```

```
</ud-speaker-information>
```

```
</speaker>
```

Ab hier der zweite Sprecher:

```
<speaker id="SPK2">
```

```
<abbreviation>S2</abbreviation>
```

```
<sex value="m"/>
```

```
<languages-used><language lang="de"/></languages-used>
```

```
<l1><language lang="de"/></l1>
```

```
<l2></l2>
```

```
<ud-speaker-information>
```

```
<ud-information attribute-name="Alter">33</ud-information>
```

```
<ud-information attribute-name="Geburtsort">Dortmund</ud-information>
```

```
</ud-speaker-information>
```

Auch zu jedem Sprecher läßt sich ein Kommentar notieren:

```
<comment>S2 ist erkältet und spricht deshalb undeutlich.</comment>
```

```
</speaker>
```

```
</speakertable>
```

Nun ist der Transkriptionskopf komplett...

```
</head>
```

... und das eigentliche Transkript, der Rumpf, kann beginnen:

```
<basic-body>
```

Die Zeitachse wird an den Anfang gestellt:

```
<common-timeline>
```

Sie besteht aus einer Liste von Zeitpunkten. Jeder Zeitpunkt hat die Attribute "id" und "time". Die Zeit kann mit "echten" Sekunden notiert werden:

```
<tli id="T0" time="240.0"/>
```

```
<tli id="T1" time="243.7"/>
```

```
<tli id="T2" time="244.0"/>
```

```
<tli id="T3" time="244.5"/>
```

```
<tli id="T4" time="245.2"/>
```

```
<tli id="T5" time="245.6"/>
```

```
<tli id="T6" time="247.2"/>
```

```
<tli id="T7" time="247.9"/>
```

```
<tli id="T8" time="248.4"/>
```

```
<tli id="T9" time="249.0"/>
```

```
</common-timeline>
```

Ab hier die Transkriptzeilen mit jeweils vier Attributen.

Hier die erste Zeile, sie ist durch das "id"-Attribut identifizierbar. Sie gehört zum Sprecher mit dem internen "id"-Attribut des Inhalts "SPK1" und fällt in die Kategorie "v"erbale, soll also das verschriftlichte Gesprochene aufnehmen:

```
<tier id="TIE1" speaker="SPK1" category="v" type="t">
```

Innerhalb der Zeile gibt es Ereignisse mit Start- und Ende-Kennzeichnungen, die über die Zeitachse eventuell mit tatsächlichen Zeitpunkten verknüpft werden können. Diese Ereignisse stellen die Zellen der Transkriptzeilen dar:

```
<event start="T0" end="T1">Kommst du morgen mit uns mit?</event>
```

```
<event start="T3" end="T4">zum Sch</event>
```

```
<event start="T4" end="T5">wimmen.</event>
```

Die erste Transkriptzeile ist nun komplett und wird abgeschlossen:

```
</tier>
```

Die zweite Transkriptzeile dient der "i"ntonation des "v"erbales, ist also für "a"nnotationen vorgesehen (aber leer):

```
<tier id="TIE2" speaker="SPK1" category="nv" type="a">
```

```
</tier>
```

Hier "v"erbales vom zweiten Sprecher:

```
<tier id="TIE0" speaker="SPK2" category="v" type="t">
```

```
<event start="T1" end="T2">hm` </event>
```

```
<event start="T2" end="T3"> ((0,5s)) </event>
```

```
<event start="T4" end="T5">mór</event>
```

```
<event start="T5" end="T6">gen lau/ äh </event>
```

```
<event start="T6" end="T7">fähr </event>
```

```
<event start="T7" end="T8">ich </event>
```

```
<event start="T8" end="T9">in die Stadt. </event>
```

Die zweite Transkriptzeile dient der "i"ntonation des "v"erbales, ist also für "a"nnotationen vorgesehen (aber leer):

```
<tier id="TIE3" speaker="SPK2" category="i" type="a">
```

```
<event start="T6" end="T7">---- </event>
```

```
</tier>
```

Hier eine Zeile für "n"on-"v"erbales:

```
<tier id="TIE4" speaker="SPK2" category="nv" type="a">
```

```
<event start="T6" end="T8">zeigt mit dem Daumen zum Fenster </event>
```

```
</tier>
```

Und so wird eine "wav"-formatige "aud"iodatei an das komplette Transkript (von "T0" bis "T9") gebunden:

```
<tier id="TIE5" category="wav" type="l">
```

```
<event start="T0" end="T9" medium="aud"
```

```
url="D:\EXMARALDA\ANNODEMO.WAV">Tonaufnahme</event>
```

```
</tier>
```

Das Transkript ist erst vollständig, wenn auch die beiden noch offenen Klammern wieder geschlossen wurden:

```
</basic-body>
```

```
</basic-transcription>
```

Ende der XML-Datei.

Wolfgang Schneider
Info@Ton-und-Text.de

Veröffentlicht am 23.2.2003

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.