

# EXMARaLDA - ein Modellierungs- und Visualisierungsverfahren für die computergestützte Transkription gesprochener Sprache

Thomas Schmidt  
SFB 538 "Mehrsprachigkeit"  
Max Brauer Allee 60  
22765 Hamburg, Germany  
thomas.schmidt@uni-hamburg.de

## Abstract

This paper attempts a new look at computer assisted transcription as it is commonly practised within the fields of discourse analysis and language acquisition studies. The first part proposes a bridge between discourse analytical methodology and text technological methods with the concept of modelling as its central idea. The second part demonstrates the EXMARaLDA system, a set of formats and tools for computer assisted transcription that builds on the ideas developed in the first part and implements them in a way that can lead to significant improvement in current research practice.

## Einleitung

Die Transkription spontansprachlicher Aufnahmen ist seit vielen Jahren ein integraler methodischer Bestandteil der Gesprächsanalyse und Spracherwerbsforschung<sup>1</sup>. Der Einsatz computergestützter Verfahren ist dabei mittlerweile die Regel, wurde bislang aber kaum eigens methodologisch reflektiert. In der wissenschaftlichen Praxis hat diese mangelnde Reflexion dazu geführt, dass diverse Verfahren zur computergestützten Transkription<sup>2</sup> scheinbar unverbunden

---

<sup>1</sup> Die Termini „Gesprächsanalyse“ und „Spracherwerbsforschung“ werden hier als Überbegriffe für eine Vielzahl sprachwissenschaftlicher Forschungsrichtungen verwendet. Stellvertretend für erstere sei die Diskursanalyse Ehlich/Rehbeinscher Prägung, für letztere die Studien zum L1-Erwerb im Rahmen der Generativen Syntax (z.B. Meisel) genannt.

<sup>2</sup> Wiederum stellvertretend seien syncWriter (Rehbein et al. 1993) und HIAT-DOS (Ehlich 1992) als Beispiele für Verfahren der computergestützten dis-

nebeneinander stehen, was einen Datenaustausch, die Erstellung größerer Korpora und die Weiterentwicklung von Datenformaten und Werkzeugen wesentlich erschwert. In ersten Teil dieses Beitrages wird argumentiert, dass über eine methodologisch motivierte Zusammenführung des etablierten Transkriptionsverständnisses mit aktueller texttechnologischer Forschung eine substantielle Verbesserung computergestützter Transkriptionsverfahren erzielt werden kann. Das im zweiten Teil vorgestellte System EXMARaLDA demonstriert, wie diese Verbesserungen in eine praxisfähige Implementierung umgesetzt werden.

## 1 Transkription als Modellierung und Visualisierung gesprochener Sprache

### 1.1 Etabliertes Transkriptionsverständnis

Das Transkriptionsverständnis, das auch heute noch als vorherrschend in der Praxis gesprächsanalytischer Forschung angesehen werden kann, hat sich bereits vor der flächendeckenden Verbreitung computergestützter Verfahren herausgebildet. Es lässt sich anhand zweier zentraler Begriffe charakterisieren:

Zum einen wird der Begriff der *Transkription als Verschriftlichung* verwendet, um den „Transfer“ (Redder 2002) oder auch „medialen Übergang“ (Deppermann 1999) von der Mündlichkeit zur Schriftlichkeit, der beim Transkribieren vollzogen wird, zu beschreiben.

Zum anderen wird mit dem Begriff der

---

kursanalytischen Transkription und CHAT (MacWhinney 2000) als ein Beispiel für ein verbreitetes Verfahren im Rahmen der Spracherwerbsforschung genannt.

*Transkription als Theorie* (Ochs 1979) die Tatsache problematisiert, dass gesprochenes Original und „geschriebenes“ Transkript keine isomorphen Abbilder voneinander sein können, sondern dass jeder Transkription ein unvermeidlicher Selektions- und Interpretationsprozess innewohnt, der unmittelbar von der zugrunde gelegten linguistischen Theorie und den Forschungsinteressen der Transkribierenden abhängt.

Bei der Bewertung von bestehenden Transkriptionssystemen schlagen sich diese beiden Aspekte in einander zuwider laufenden Gütekriterien nieder: Kriterien der *Lesbarkeit* nehmen Bezug auf die Transkription als Verschriftlichung und fordern, dass bezüglich der verwendeten Zeichen und ihrer räumlichen Anordnung möglichst ikonisch, ökonomisch und „schriftnah“ verfahren wird (z.B. DuBois 1991). Dem gegenüber verlangen Kriterien der *Beschreibungsadäquatheit* eine größtmögliche Explizitheit und Präzision der verwendeten Kategorien sowie eine auf das Unvermeidliche reduzierte „schriftsprachliche Korrektur“ des Gesprochenen – beides sind Anforderungen, die nur mit Rückgriff auf den Aspekt der Transkription als Theorie zu verstehen sind, und die mit den Kriterien der Lesbarkeit konkurrieren.

Zusammenfassend lässt sich das etablierte Transkriptionsverständnis in der Gesprächsanalyse und Spracherwerbsforschung also unter das Schlagwort der „theoriegeleiteten Verschriftlichung“ fassen. Weil die daraus resultierenden widersprüchlichen Anforderungen an das Transkript in der Literatur zwar thematisiert, nicht jedoch aufgelöst werden, verstehen sich bestehende Transkriptionssysteme als Kompromisse zwischen Lesbarkeit und Beschreibungsadäquatheit. Die Möglichkeiten, die sich speziell durch den Computereinsatz in dieser Hinsicht ergeben, haben bislang keinen Eingang in die breitere wissenschaftliche Diskussion gefunden – das etablierte Transkriptionsverständnis bezieht den Computer nicht als wesentliche Komponente in seine Betrachtungen ein.

## **1.2 Der Modellbegriff in der Informatik und Texttechnologie**

Der Modellbegriff spielt sowohl in der Informatik im allgemeinen Sinne, als auch im spezialisierten Zweig der Texttechnologie eine wesentliche Rolle bei der Charakterisierung der disziplin

lineigenen Methoden. In diesem Zusammenhang ist zunächst der von Chen (1976) eingeführte Begriff des *Entity-Relationship-Modelling* zu nennen, der jegliche Datenverarbeitung auf dem Rechner als ein Anfertigen und Manipulieren symbolischer Modelle begreift. Wie z.B. Date (1995) darlegt, stellt diese Sichtweise das vorherrschende Paradigma der Datenbanktheorie dar, auf das sich alle spezifischeren Datenmodelle (z.B. das relationale) gründen.

Innerhalb der Texttechnologie verwendet Lobin (2000) weiterhin den Terminus der *Informationsmodellierung*, um die Methoden der computergestützten Beschreibung und Verarbeitung speziell sprachlicher Daten zu charakterisieren. Darauf aufbauend lässt sich die für die Texttechnologie zentrale Unterscheidung zwischen logischer und graphischer Struktur von Dokumenten ebenfalls auf modelltheoretische Konzepte zurückführen: derjenige Aspekt eines Dokumentes, der einer rechnergestützten Manipulation in Form vom Markup o.Ä. zugänglich gemacht wird, stellt eine *rein symbolische* Modellierung der Dokumentinformation dar. Für die Rezeption durch den menschlichen Benutzer wird diese logische Struktur hingegen in der Regel in eine teilweise *analoge* (oder auch ikonische) Form – eine Anzeige am Bildschirm oder auf dem Papier – überführt. Der Mehrwert texttechnologischer Verfahren ergibt sich demnach daraus, dass sie einerseits eine rein symbolische Modellierung von Dokumentinformation ermöglichen, die sich aufgrund ihres höheren Abstraktionsgrades (vgl. auch Stachowiak 1973) besser für eine formale rechnergestützte Manipulation eignet, dass sie andererseits aber gleichzeitig Methoden bereitstellen, um symbolische in analoge Modelle zu überführen, die einer Verarbeitung durch den menschlichen Benutzer eher entgegen kommen.

## **1.3 Transkription als Modellierung und Visualisierung**

Der in Schmidt (i.V.) ausgearbeitete Ansatz besteht nun darin, mit Hilfe des Modellbegriffes das in der Gesprächs- und Spracherwerbsforschung etablierte Transkriptionsverständnis mit den Methoden der Informatik und Texttechnologie zusammenzuführen. Damit wird zum einen das Ziel eines besseren methodologischen Verständnisses des Computereinsatzes bei der Transkription verfolgt, zum anderen soll auf

diese Weise aktuelle texttechnologische Forschung unmittelbar für die Praxis der Gesprächs- und Spracherwerbsforschung nutzbar gemacht werden. Zwei Argumente sind dabei zentral:

Erstens lässt sich die traditionell als Theorieabhängigkeit bezeichnete Eigenschaft von Transkriptionen ohne Weiteres auch als Modellhaftigkeit begreifen – der Stellvertretungscharakter einer Transkription und die ihr innewohnende Selektivität und Zweckgebundenheit beschreiben genau diejenigen Merkmale (Abbildungs-, Verkürzungs- und Pragmatisches Merkmal), die z.B. Stachowiak (1973) als allgemeingültige Kennzeichen wissenschaftlicher Modellbildung nennt. Eine Brücke zwischen gesprächsanalytischen und texttechnologischen Methoden kann somit zunächst durch einen einfachen Perspektivwechsel terminologischer Natur geschlagen werden, indem Transkription statt als „theoriegeleitete Verschriftlichung“ als eine „Modellbildung mit den Mitteln der Schrift“ verstanden wird.

Zweitens zeigt die texttechnologische Unterscheidung zwischen logischer und graphischer Struktur von Dokumenten Auswege aus dem traditionellen Dilemma der Transkription auf: Der Computer als Transkriptioninstrument macht es praktikabel, das Kriterium der Beschreibungsadäquatheit allein auf die rein symbolische Repräsentation anzuwenden und das Kriterium der Lesbarkeit auf die analoge graphische Darstellung zu beschränken. Wenn computergestützte Transkription in diesem Sinne als ein zweiseitiger Prozess der *Modellierung und Visualisierung* aufgefasst wird, müssen Transkriptionssysteme nicht mehr als unvermeidbarer Kompromiss zwischen widersprüchlichen Anforderungen ausgehandelt werden.

## 2 Modellierungsverfahren

Wenn – wie oben dargelegt – Transkriptionen Modelle gesprochener Sprache sind, deren Verschiedenheit sich aus dem theoretischen Hintergrund und den Forschungsinteressen ihrer Nutzer motiviert, so ist die Diversität existierender Transkriptionssysteme weniger ein beklagenswerter Mangel denn eine wissenschaftliche Notwendigkeit. Der Versuch einer „Standardisierung“ verschiedener Transkriptionssysteme erscheint vor dieser Überlegung wenig sinnvoll<sup>3</sup>.

<sup>3</sup> Dennoch wird sie immer wieder versucht, vgl. z.B.

Vielmehr kann eine Verbesserung hinsichtlich der Austauschbarkeit von Transkriptionsdaten nur erreicht werden, indem über Spezifika einzelner Modelle (i.e. Transkriptionssysteme) abstrahiert und nach einem *Modellierungsverfahren* gesucht wird, mittels dessen sich die einzelnen Modelle auf einer gemeinsamen formalen und strukturellen Basis ausdrücken lassen. Mit anderen Worten: zu fragen ist nicht nach *einer* geeigneten Beschreibungsmethode für Daten gesprochene Sprache, sondern danach, wie – eine Abstraktionsstufe höher – solche Beschreibungen prinzipiell zu organisieren sind, damit sie einer computergestützten Verarbeitung möglichst gut zugänglich werden. Zwei solcher Modellierungsverfahren sind in den letzten Jahren vorgeschlagen worden:

Das Konzept der *Standoff-Annotation* bzw. des NITE Object Models (Carletta et al. 2002, Evert et al. 2003) beschreibt Transkriptionsdaten als ein System überlappender *Hierarchien*. Der praktische Vorteil dieser Vorgehensweise ergibt sich aus ihrer Nähe zum OHCO-Paradigma (De Rose et al. 1990) und dessen grundlegender Bedeutung für die Architektur von Markup-Sprachen: wenn Hierarchien als die primären Relationen der Datenbeschreibung angesehen werden, lassen sich Sprachen wie SGML und XML und die zugehörigen Technologien (z.B. XSLT-Stylesheets oder XPath-Queries) einfach und effektiv anwenden. Wie in Schmidt (i.V.) ausgeführt, zeichnen sich aber gerade Transkriptionen gesprochener Sprache durch eine Vielzahl von Relationen aus, die sich nicht ohne Weiteres in eine oder mehrere Hierarchien einfügen. Insbesondere wird die für viele Untersuchungszwecke bedeutsame Relation der zeitlichen Parallelität zwischen Gesprächseinheiten (mehrer Sprecherbeiträge oder Modalitäten überlappen einander) im Konzept der Standoff-Annotation stark vernachlässigt.

Demgegenüber werden im Konzept der *Annotationsgraphen* (Bird/Lieberman 2001) Transkriptionsdaten als gerichtete azyklische *Graphen* (DAGs) beschrieben, deren Knoten den immer vorhandenen Zeitbezug von Gesprächseinheiten repräsentieren und deren Kanten die nicht-zeitliche Information tragen. Wie die Autoren zeigen, lassen sich in dieser Weise sowohl zeit-

---

Barnett et al. (2000) oder auch die TEI-Konventionen.

liche Parallelität und Sequenzialität als auch andere strukturelle Beziehungen wie Äquivalenz oder Dominanz in einer vereinheitlichten Weise darstellen. Der wichtigste Einwand gegen diesen Ansatz betrifft seine rechnerische Komplexität: die Annahme einer sehr allgemeinen Struktur wie DAGs macht es schwierig, effektive Verarbeitungsalgorithmen (z.B. zur Suche oder Visualisierung) zu entwerfen, und – anders als im Falle der überlappenden Hierarchien der Standoff-Annotation – steht keine weit verbreitete, standardisierte Technologie zur Verfügung, die diesem Datenverständnis direkt entgegenkommen würde.

Das in Schmidt (i.V.) entwickelte Modellierungsverfahren der *Transkriptionsgraphen* baut auf dieser Kritik auf und entwirft eine strukturell eingeschränkte und semantisch spezifizierte Version von Annotationsgraphen, die einerseits noch allgemein genug ist, um verschiedene bestehende Transkriptionssysteme ausdrücken zu können, andererseits aber auch einfach genug, um Basis einer in der Praxis einsatzfähigen Implementierung werden zu können. Das im folgenden Abschnitt vorgestellte EXMARaLDA-System beinhaltet wesentliche Elemente einer solchen Implementierung.

### 3 EXMARaLDA

EXMARaLDA steht für „Extensible Markup Language for Discourse Annotation“ und wird am Sonderforschungsbereich „Mehrsprachigkeit“ der Universität Hamburg entwickelt. Es fungiert als zentrale Architekturkomponente einer in der Entwicklung begriffenen „Datenbank Mehrsprachigkeit“, die als gemeinsame Plattform aller SFB-Projekte zur Erstellung, zum Austausch und zur Auswertung von Korpora gesprochener Sprache genutzt werden soll.<sup>4</sup> Die Ausgangslage entspricht dabei der oben geschilderten: abhängig von ihrem theoretischen Hintergrund und ihren Untersuchungsinteressen verwenden die einzelnen Projekte unterschiedliche Transkriptionssysteme und erstellen und bearbeiten ihre Daten jeweils mit eigens für „ihr“ System konzipierten Software-Werkzeugen. Ein Datenaustausch zwischen Projekten, eine Zusammenlegung von Korpora oder

gar eine Weiterverarbeitung von Daten eines Projektes mit den Werkzeugen eines anderen werden so schwierig bis unmöglich. Indem EXMARaLDA sich als Modellierungsverfahren im oben beschriebenen Sinne begreift, versucht es, diese Schwierigkeiten zu überwinden.

#### 3.1 Datenformate

EXMARaLDA übernimmt das Konzept der Annotationsgraphen, den immer vorhandenen Zeitbezug von Gesprächseinheiten als primäres strukturelles Ordnungsprinzip für die Datenbeschreibung zu verwenden. Somit lassen sich auch EXMARaLDA-Daten immer als gerichtete azyklische Graphen auffassen. Aus den oben genannten Gründen werden jedoch zusätzliche strukturelle Einschränkungen ans diesen Graphen vorgenommen:

In der einfachsten Version, die sich im Datenformat der *Basic-Transcription*<sup>5</sup> wieder findet, wird gefordert, dass alle Einheiten sich zum einen auf eine einzige vollständig geordnete Zeitachse einordnen lassen. Zum anderen muss eine Partition der Einheiten in Spuren existieren, so dass Einheiten innerhalb einer Spur einander nicht überlappen. Zusätzlich muss jeder dieser Spuren ein Element aus einer endlichen Menge von Sprechern und Kategorien zugeordnet werden können:

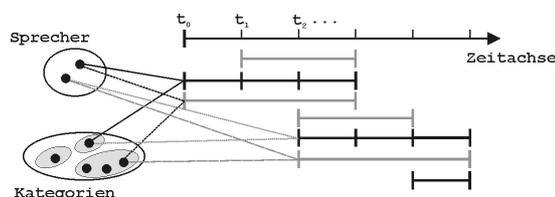


Abb. 1: Zeitachse und Spuren einer *Basic-Transcription*

Für viele Zwecke der Verarbeitung ist diese stark vereinfachte Form bereits ausreichend – beispielsweise lässt sich eine *Basic-Transcription* als Transkription in Partiturnotation eingeben und visualisieren (s.u.), und auch für andere gebräuchliche Notationsformen stellen die so kodierten Informationen eine hinlängliche Basis dar. Für einige weiterreichende Verarbeitungs-

<sup>4</sup> Darüber hinaus werden die EXMARaLDA-Formate und -Werkzeuge mittlerweile aber auch außerhalb des SFB in Forschung und Lehre eingesetzt.

<sup>5</sup> Physikalisch werden *Basic-Transcription* und *Segmented-Transcription* als XML-Dateien gespeichert, ohne jedoch dabei dem XML inhärenten hierarchischen Datenmodell zu folgen (s.o.).

formen – insbesondere fortgeschrittenere Annotationen auf der Wort- oder Silbenebene – erweist sich die Beschränkung auf eine einzige, vollständig geordnete Zeitachse jedoch als zu stark. EXMARaLDA definiert daher mit der *Segmented-Transcription* ein zweites Datenformat, das Verzweigungen der Zeitachse zulässt. Transkriptionseinheiten, für die ein Transkribent keine eindeutige zeitliche Ordnung feststellen kann (in der Abb. z.B. die Wörter „du“ und „sage“), lassen sich so dennoch in das zeitbasierte Datenmodell integrieren:

A: das [sagst du so]  
 B: [das sage ] ich

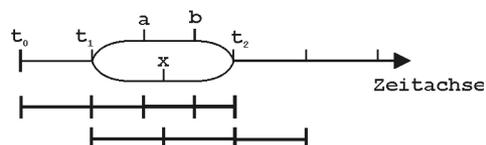


Abb. 2 : Verzweigende Zeitachse einer *Segmented-Transcription*. Für die Zeitpunkte a und x muss keine Ordnung festgelegt werden.

### 3.2 Parametrisierbare Segmentierung

Das EXMARaLDA-System sieht vor, dass Eingabe-Werkzeuge in erster Linie auf dem einfacheren (und damit benutzerfreundlicheren) *Basic-Transcription*-Format operieren. Damit das komplexere *Segmented-Transcription*-Format und die damit verbundenen fortgeschrittenen Annotations- und Auswertungsmöglichkeiten genutzt werden können, müssen *Basic-Transcriptions* nach Abschluss der Eingabe automatisch in *Segmented-Transcriptions* überführt werden. Dies geschieht mittels Endlicher Maschinen (Finite State Machines), die implizites Markup (z.B. Interpunktionszeichen) in explizites Markup umwandeln<sup>6</sup>. So würde z.B. eine geeignete Endliche Maschine folgende, mit den HIAT-Konventionen (Rehbein et al. 1993) konforme Zeichenkette in Äußerungen (U), Wörter

<sup>6</sup> Dieser Vorgang ist vergleichbar, aber nicht identisch mit der Tokenisierung geschriebener Texte. Der wesentliche Unterschied liegt darin, dass bei Transkriptionen die Interpunktion als Grundlage der Segmentierung vollständig vorhersagbar ist, weil sie durch formale Transkriptionskonventionen geregelt wird, während bei einer Tokenisierung anhand standardorthographischer Interpunktionsregeln Mehrdeutigkeiten auftreten können.

(w) und nicht-phonologische Einheiten (NP) segmentieren:

Basic-Transcription:

Ja ((hustet)) sicher. Oder?

Endliche Maschine:



HIAT-FSM

Segmented-Transcription:



	U		U
W	NP	W	W
Ja	hustet	sicher	oder

Abb. 3 : Segmentierung mittels Endlicher Maschinen

Da dieser Prozess in offensichtlicher Weise von der verwendeten Transkriptionskonvention abhängig ist (denn diese regelt, welche zu segmentierenden Einheiten durch welches implizite Markup kenntlich gemacht sind), ermöglicht das EXMARaLDA-System die Verwendung verschiedener Endlicher Maschinen zum Segmentieren.<sup>7</sup> Auf diese Weise wird die Segmentierung parametrisierbar gemacht, das transkriptionssystemübergreifende Modellierungsverfahren selbst aber unberührt gelassen.

### 3.3 Eingabewerkzeuge

Als ein zentrales Bearbeitungswerkzeug für EXMARaLDA-Transkriptionen im *Basic-Transcription*-Format wurde der EXMARaLDA Partitur-Editor entwickelt. Dieser ermöglicht u.a. das Eingeben und Editieren einer Transkription in Partitur-Notation, das Verknüpfen von Teilen einer Transkription mit externen Bild-, Ton- oder Videodateien, die parametrisierbare Ausgabe von Partituren in verschiedene Präsentationsformate (z.B. RTF und HTML, s.u.) und die Verwaltung von Sprecherinformationen und Metadaten zur Transkription. Zusätzlich fungiert er als graphische Benutzerschnittstelle für diverse Import- und Exportfilter (s.u.) sowie für die

<sup>7</sup> Derzeit existieren Endliche Maschinen für die wichtigsten deutschsprachigen gesprächsanalytischen Transkriptionssysteme HIAT, GAT und DIDA sowie für das in der Spracherwerbsforschung verbreitete System CHAT.

automatische Segmentierung der Transkriptionsdaten (s.o.).

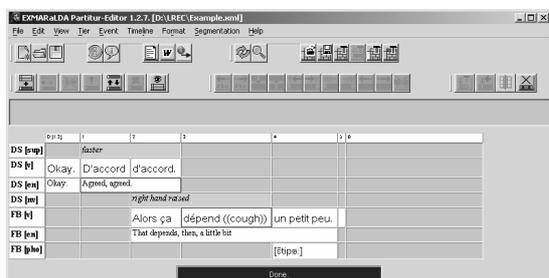


Abb. 4 : EXMARaLDA Partitur-Editor

Entscheidend für das Konzept von EXMARaLDA ist jedoch, dass die Eingabe und das Bearbeiten von Transkriptionen gerade *nicht* fest an eine bestimmte Software gebunden sind. Weil die wesentlichen Merkmale einer EXMARaLDA *Basic-Transcription* – eine vollständig geordnete Zeitachse und die Partition der Transkriptionseinheiten in Spuren – auch für mehrere andere derzeit gebräuchliche Datenformate Gültigkeit besitzen, können über einfache Import- und Exportfilter auch mit anderen Werkzeugen erstellte Daten für die EXMARaLDA-Architektur nutzbar gemacht werden. Drei solcher Werkzeuge sind dabei von besonderem Interesse:

- das phonetische Analyseprogramm Praat von der Universität Amsterdam,
- das Video-Annotationsprogramm TASX-Annotator (Milde/Gut 2001) von der Universität Bielefeld und
- die ELAN-Software (Brugman 2003) des MPI in Nijmegen.

Weiterhin ermöglicht der „Simple EXMARaLDA“-Importfilter die Überführung von Transkriptionen aus einfachen Textdateien in das EXMARaLDA-Format und gestattet so zusätzlich die Verwendung von Texteditoren und Textverarbeitungsprogrammen als Eingabewerkzeuge.

### 3.4 Visualisierung

Die Möglichkeit, mit dem Computer erstellte Transkriptionen für eine eingehende Analyse auf Papier oder Bildschirm visualisieren zu können, ist ein Aspekt, der zwar von entscheidender Bedeutung für die gesprächsanalytische Arbeitsweise ist, der jedoch von vielen derzeit in der Entwicklung befindlichen Werkzeugen ver-

nachlässigt wird<sup>8</sup>. Insbesondere wird die Tatsache, dass die Anforderungen an ein computergerechtes Datenformat oft den Anforderungen an eine für den menschlichen Benutzer geeignete Visualisierung direkt zuwiderlaufen, kaum thematisiert. Wie bereits erwähnt, finden die oben erörterten „klassischen“ Gegensätze zwischen Beschreibungsadäquatheit und Lesbarkeit von Transkripten hier ihre „moderne“ Entsprechung, dies jedoch mit dem Unterschied, dass sie sich auf dem Rechner mittels moderner texttechnologischer Verfahrensweisen überwinden lassen.

In EXMARaLDA wird daher besonderer Wert darauf gelegt, dem (für die computergestützte Verarbeitung konzipierten) Modellierungsverfahren geeignete (für den menschlichen Betrachter konzipierte) Visualisierungsmethoden an die Seite zu stellen. Das technisch anspruchsvollste Problem ist dabei die Ausgabe von Transkriptionen als umgebrochene Partituren (vgl. auch Bow et al. 2003). EXMARaLDA ermöglicht eine solche Ausgabe in mehreren Präsentationsformaten (RTF, HTML, PDF), wobei eine Vielzahl von Parametern (Schriftarten und -größen, Seitenformate etc.) durch den Benutzer einstellbar sind:

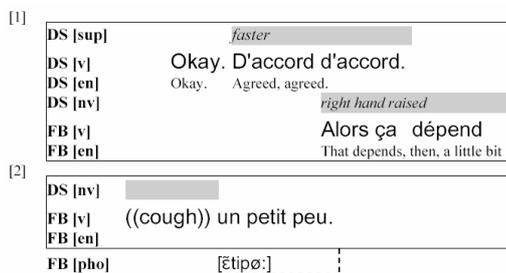


Abb. 5 : Visualisierung als umgebrochene Partitur

Über geeignete XSLT-Styleheets (und teilweise im Zusammenspiel mit den Segmentierungsfunktionen) ermöglicht EXMARaLDA jedoch darüber hinaus eine Vielzahl weiterer Visualisierungsformen. Stellvertretend seien die klassischerweise mit der Partiturnotation konkurrierenden Notationsverfahren der Spalten- und Zeilennotation genannt.

<sup>8</sup> So enthalten auch die oben erwähnten Konzepte der Standoff-Annotation und der Annotationsgraphen kaum Überlegungen zur benutzergerechten Visualisierung der jeweiligen Datenmodelle.

DS [right hand raised]  
 Okay. D'accord d'accord.  
 {faster  
 Okay. | Agreed, agreed. }  
 FB Alors ça dépend ((cough)) un petit peu.  
 {That depends, then, a little bit  
 [ɛ̃tipø:]}

DS [sup]	DS [v]	DS [en]	DS [nv]	FB [v]	FB [en]	FB [pho]
0	Okay.	Okay.				
1	faster	D'accord		Alors ça		
2		d'accord.	right hand raised	(dépend ((cough)))	That depends, then, a little bit	
3				un petit peu.		[ɛ̃tipø:]
4						
5						
6						

Abb. 6 : Visualisierung in Zeilen- und Spaltennotation

### 3.5 Auswertung

Der wahrscheinlich wichtigste Mehrwert computergestützter Transkriptionsverfahren gegenüber dem traditionellen Transkribieren mit Bleistift und Papier liegt in der Möglichkeit, die Auswertung der Daten – insbesondere die Suche und Quantifizierung spezifischer Gesprächsphänomene – teilautomatisiert vornehmen zu können. Durch die in EXMARaLDA praktizierte Trennung zwischen computergerechtem symbolischem Modell und benutzerorientierter analoger Visualisierung werden solche teilautomatisierten Auswertungsverfahren wesentlich einfacher als in den bis dato existierenden Systemen umsetzbar. Wiederum stellvertretend seien zwei Instrumente genannt, die sich bereits in der Praxis der täglichen Arbeit am Sonderforschungsbereich bewährt haben:

Zum einen ermöglicht ein Konkordanzwerkzeug die Erstellung von KWIC<sup>9</sup>-Konkordanzen, mittels derer Gesprächsphänomene systematisch gesucht und in ihrem unmittelbaren Kontext betrachtet und analysiert werden können.

Sprecher (TL)	Vorher	Nachher
TE [41]	S	Tvegginnú
TE [47]	E	ginnú? (líkist 0,5s)
IM [002]	S	inn
SR [27]	E	handla on
IN [0,50]	S	Donn
IN [0,77]	S	Þú hefurðu skilið?
IN [0,80]	S	meðan
IN [0,84]	S	enligt
IN [0,91]	S	vandfr
IN [0,94]	S	jag
IN [0,97]	S	vet
IN [0,99]	S	myndet

Abb. 7: KWIC-Konkordanz

Zum anderen können Transkriptionen, die mittels der oben beschriebenen Methode in sprachliche Einheiten segmentiert wurden, in RDB-

<sup>9</sup> Key Word In Context

Systeme (im Beispiel MS Access) überführt werden und dort um analytische Annotationen ergänzt und mittels SQL-Queries durchsucht und ausgewertet werden.

SPRECHER	WORT	AUSSERUNG
S	mycket	Den som gapar efter mycket...
S	lite	Jag ska andas in här lite
S	lite	Åh ja, jag ringar runt och visar lite vad jag kan.
S	lite	Ja, det blir något av det, och jag var lite orolig för att hela affären påminner ge
S	mycket	Åh, har ni någon som är fakalieneurotiker som har alltså väldigt mycket prob
S	mycket	Och/och sen det sista är...vi hade ett problem här då första säsongen att c
S	mycket	Ja, det blir något av det, och jag var lite orolig för att hela affären påminner ge
N	litt	Man sa jo att dere ringte, og så ble vi litt nervøse
N	litt	Jeg tror det har litt med det å gjøre
N	mye	Og så...dyppet seentallene og gikk nedover, og nå går de oppover igjen, men

Abb. 8: RDB-Abfrage

Weitere Auswertungswerkzeuge befinden sich in Planung.

### 3.6 Anwendungen

EXMARaLDA und der Partitur-Editor haben mittlerweile das Beta-Stadium hinter sich gelassen und werden in zahlreichen Projekten in Forschung und Lehre eingesetzt. Stellvertretend seien hier einige Projekte am Sonderforschungsbereich „Mehrsprachigkeit“ genannt, die EXMARaLDA für ihre Korpusarbeit einsetzen:

- Das Projekt-Korpus „Semikommunikation und rezeptive Mehrsprachigkeit im heutigen Skandinavien“ setzt sich aus Aufnahmen von Schulstunden, Radiosendungen und Diskussionen zusammen, in denen Sprecher des Schwedischen, Dänischen und Norwegischen in ihrer jeweiligen Muttersprache miteinander kommunizieren. Die ursprünglich mit HIAT-DOS angefertigten Transkriptionen wurden nach EXMARaLDA überführt, neue Transkriptionen werden mit Praat in Kombination mit dem EXMARaLDA Partitur-Editor erstellt.
- Das Projekt „Sprachliche Konnektivität bei bilingual türkisch-deutsch aufwachsenden Kindern“ vergleicht den Spracherwerb bei monolingualen und bilingualen Sprechern des Türkischen. Das Korpus besteht aus zahlreichen Aufnahmen freier Kommunikationen, die teilweise durch ein vereinheitlichtes Aufnahme-Setting (sog. Evokative Feldexperimente) besser vergleichbar gemacht werden. Auch hier lagen ältere Datenbestände (im syncWriter-Format) vor, die nach EXMARaLDA überführt wurden, wäh-

rend die Erstellung neuer Daten mittlerweile direkt mit dem EXMARaLDA Partitur-Editor erfolgt.

- Das Projekt „Prosodische Beschränkungen zur phonologischen und morphologischen Entwicklung im bilingualen Spracherwerb“ befasst sich mit phonologisch/phonetischen Aspekten des Spracherwerbs bei spanisch/deutsch bilingualen Kindern. Auch hier wurden ältere WordBase-Datenbestände zunächst nach EXMARaLDA überführt und anschließend die Eingabemethode auf das neue System umgestellt.

Die Gesamtheit dieser (und weiterer) Korpora zeigt deutlich, dass der Einsatz texttechnologischer Verfahren und ein verändertes Verständnis computergestützter Transkription tatsächlich zu konkreten Verbesserungen in der Praxis der Gesprächs- und Spracherwerbsforschung beitragen können: Transkriptionsdaten, die in verschiedenen Forschungszusammenhängen mit unterschiedlichsten Zielsetzungen erstellt wurden, werden durch den Einsatz eines generischen Modellierungsverfahrens dennoch bis zu einem gewissen Grade vergleichbar, und Software-Lösungen lassen sich über spezifische Projektzusammenhänge hinweg verwenden.

## Literatur

- Barnett, R. et al. (2000): *The LIDES Coding Manual*. International Journal of Bilingualism, 131-270.
- Bird, S. and Liberman, M. (2001): *A formal framework for linguistic annotation*. Speech Communication 33(1,2), 23-60.
- Bow, C. et al. (2003): *Towards a General Model Of Interlinear Text*. In “Proceedings of the E-Meld Workshop”. LSA Institute, Michigan State University, Lansing.
- Brugman, Hennie (2003): *Annotated Recordings and Texts in the DoBeS Project*. In “Proceedings of the E-Meld Workshop”. LSA Institute, Michigan State University, Lansing.
- Carletta, J. et al. (2002): *Supporting linguistic annotation using XML and stylesheets*. <http://www.cogsci.ed.ac.uk/~jeanc/revised.but.like.readings-in-corpling.pdf>
- Chen, P. (1976): *The Entity-Relationship Model – Toward a Unified View of Data*. In “Readings in Database Systems”, M. Stonebraker, ed., Morgan Kaufmann Publishers, San Metao, pp.374-387.
- Date, C. (1995): *An Introduction to Database Systems*. Addison-Wesley, New York.
- Deppermann, A. (1999): *Gespräche analysieren: eine Einführung*. Leske & Budrich, Opladen.
- De Rose, S. et al. (1990): *What is Text, Really?* Journal of Computing in Higher Education 1(2), 3-26.
- DuBois, John (1991): *Transcription design principles for spoken discourse research*. Pragmatics 1, 71-106.
- Ehlich, Konrad (1992): *HIAT - a Transcription System for Discourse Data*. In: „Talking Data – Transcription and Coding in Discourse Research.“, J. Edwards & M. Lampert, eds., Erlbaum, Hillsdale, pp. 123-148.
- Evert, S. et al. (2003): *The NITE Object Model. Version 2.1*. NITE Internal document. <http://www.ltg.ed.ac.uk/NITE/documents.html>
- Lobin, H. (2000): *Informationsmodellierung in XML und SGML*. Springer, Berlin.
- MacWhinney, B. (2000): *The CHILDES project: tools for analyzing talk*. Lawrence Erlbaum, Mahwah, NJ.
- Milde, J.-T. / Gut, U. (2001): *The TASX-Environment: an XML-based corpus database for time aligned language data*. In “Proceedings of the IRCS Workshop On Linguistic Databases” S. Bird et al., eds., University of Pennsylvania, pp. 174-180.
- Ochs, E. (1979): *Transcription as Theory*. In: “Developmental Pragmatics.”, E. Ochs & B. Schieffelin, eds. Academic Press, London, pp. 43-72.
- Redder, A. (2002): *Professionelles Transkribieren*. In „Transkription – Medien/Lektüre“, L. Jäger & G. Stanitzek, Georg, eds., Fink, München pp. 115-131.
- Rehbein, J. et al. (1993): *Manual für das computergestützte Transkribieren mit dem Programm syncWRITER nach dem Verfahren der Halbininterpretativen Arbeitstranskriptionen (HIAT)*. Germanisches Seminar, Universität Hamburg.
- Schmidt, T. (i.V.): *Computergestützte Transkription als Modellierung und Visualisierung gesprochener Sprache mit texttechnologischen Mitteln*. Dissertation, Universität Dortmund.
- Stachowiak, H. (1973): *Allgemeine Modelltheorie*. Springer, Wien.