|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
| **How to:**  **EXMARaLDA – Basiswissen** | | |
|  | | |

Dieses Dokument erklärt die Grundkonzepte von EXMARaLDA und die Terminologie der Menüs, Dialoge und der restlichen Dokumentation.

Inhalte

[A.Transkriptionen 2](#_Toc371609505)

[1. Ereignisse, Zeitachse, Spuren und Sprecher 2](#_Toc371609506)

[2. Spurtypen und Kategorien 3](#_Toc371609507)

[3. Segmentketten 4](#_Toc371609508)

[B. Struktur der Coma-Metadaten   4](#_Toc371609509)

[1. Korpora 4](#_Toc371609510)

[2. Kommunikationen 4](#_Toc371609511)

[3. Sprecher 5](#_Toc371609512)

[4. Aufnahmen 5](#_Toc371609513)

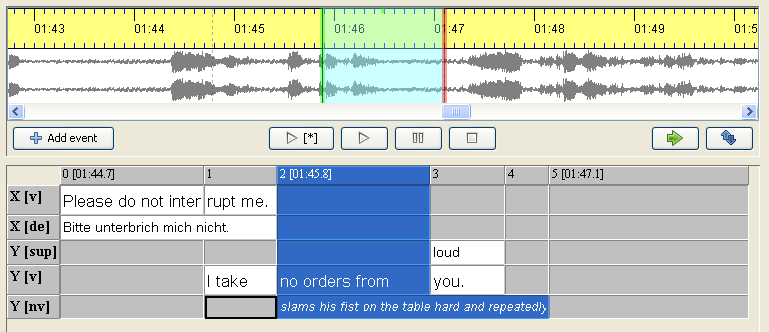
[5. Transkriptionen 5](#_Toc371609514)

[6. Weitere Dateitypen 6](#_Toc371609515)

# [Systemversion 1.5.1.]

# A.Transkriptionen

Eine EXMARaLDA-Transkription ist eine wohldefinierte Struktur, die aus einigen Grundeinheiten besteht, die miteinander in Relation gesetzt werden. Um diese Grundeinheiten der EXMARaLDA-Transkription und ihre Relationen zu erläutern, wird folgende kurze Transkription als Beispiel aufgeführt:



## Ereignisse, Zeitachse, Spuren und Sprecher

Die elementare Einheit in einer EXMARaLDA-Transkription ist das **Ereignis**. Es enthält Text, der etwas beschreibt, was in der transkribierten Aufnahme passiert.

Das obige Beispiel umfasst insgesamt acht Ereignisse – die weißen Kästchen der Partituroberfläche. Fünf davon beschreiben Wörter (oder Wortteile), die von den zwei Sprechern geäußert werden ('Please do not inter', 'rupt me. [[1]](#footnote-1)', 'I take ', 'no orders from ' und 'you. '). Ein Ereignis ('slams his fist on the table hard and repeatedly') beschreibt nonverbales Verhalten. Die übrigen zwei Ereignisse enthalten eine deutsche Übersetzung der Äußerung von Sprecher X ('Bitte unterbrich mich nicht') und suprasegmentale Merkmale ('loud') für das von Sprecher Y geäußerte Wort 'you',.

Durch den Bezug zur **Zeitachse** und zu einer **Spur** wird jedem Ereignis ein Platz in der Transkription zugewiesen.

Bei der **Zeitachse** handelt es sich um eine geordnete Abfolge von Zeitpunkten. Jeder Zeitpunkt kann mit einem absoluten Zeitwert versehen werden, der als Zeiger in die entsprechende Audio- oder Videoaufnahme interpretiert wird. Die Zeitachse im obigen Beispiel enthält insgesamt sechs Zeitpunkte – die grauen nummerierten Zellen in der obersten Zeile der Partituransicht. Drei von diesen Zeitpunkten (0, 2 und 5) sind absolute Zeitpunkte zugewiesen worden (jeweils '01:44.7', '01:45.8' und '01:47.1'). Ein Ereignis wird der Zeitachse zugeordnet, indem Start- und Endzeitpunkt auf der Zeitachse für dieses Ereignis angegeben wird. So beginnt das markierte Ereignis mit dem Text 'no orders from ' bei Zeitpunkt 2 und endet bei Zeitpunkt 3. Das Ereignis mit dem Text 'slams his fist on the table hard and repeatedly' beginnt zur gleichen Zeit, endet aber erst bei Zeitpunkt 5 und erstreckt sich somit über eine längere Zeitspanne. Da Zeitpunkt 2 ein absoluter Zeitwert zugewiesen wurde, erhalten wir zusätzlich die Information, dass ‚no orders from ‘ und ‚slams his fist on the table hard and repeatedly‘ bei '01:45:8' der transkribierten Aufnahme beginnen. Der Editor nutzt diese Information, um den entsprechenden Abschnitt im Oszillogram der Aufnahme auszuwählen.

Eine **Spur** bündelt alle Ereignisse mit ähnlichen Eigenschaften, typischerweise etwa alle Ereignisse, die eine bestimmte Art von Verhalten (z.B. verbal oder nonverbal) eines bestimmten Sprechers beschreiben. Das obige Beispiel enthält insgesamt fünf Spuren – die Zeilen nach der obersten Zeile der Partitur einschließlich der Spurlabels in den grauen Zellen ganz links (gekennzeichnet 'X [v]', 'X [de] usw.).

Jede Spur kann wiederum einem **Sprecher** zugeordnet werden. Die Sprecher werden in einer Sprechertabelle verwaltet. Die Sprechertabelle des obigen Beispiels enthält zwei Sprecher, die durch 'X' und 'Y' gekennzeichnet sind.

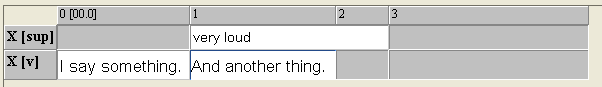
## Spurtypen und Kategorien

Neben der (optionalen) Sprecherzuordnung, wird jede Spur einem **Typ** und einer **Kategorie** zugordnet. Wenn sie Ihre Daten zuverlässig durchsuchen und umwandeln möchten, ist diese Zuordnung sehr wichtig, weil sie für viele weitere automatische Verarbeitungsschritte von Bedeutung ist.

**Kategorien** legen fest, welcher Typ von Information in den Ereignissen der jeweiligen Spur beschrieben wird. Im obigen Beispiel gibt es zwei Spuren der Kategorie 'v' mit Ereignissen, die verbales Verhalten beschreiben. Den anderen drei Spuren wurden die Kategorien 'de', 'sup' und 'nv' für 'Deutsche Übersetzung', 'suprasegmentale Informationen' und 'Nonverbales' zugeordnet. Kategorien(namen) können frei gewählt werden (d.h., man könnte genauso gut die verbale Spur mit der Kategorie 'verbal' kennzeichnen), es sollte jedoch sichergestellt werden, dass die Kategorienzuordnung für das gesamte Korpus konsistent ist. Verwenden Sie also nicht die Kategorie 'v' in einer Transkription, 'V' in der nächsten und ‚verbal‘ in einer anderen.

**Typen** sind vordefinierte Klassifikationen für Spuren. Es gibt drei unterschiedliche Typen:

* Typ **'t'** steht für 'transcription' (Transkription). Dies ist gewöhnlich der Typ für Spuren, in denen verbales Verhalten beschrieben bzw. transkribiert wird. Sie sollten nur eine einzige Spur dieses Typs für jeden Sprecher und keine ohne Sprecherzuordnung haben. Im obigen Beispiel sind die erste und die vierte Spur solchen Typs.
* Typ **'d'** steht für 'description' (Beschreibung). Dies ist gewöhnlich der Typ für Spuren, die nonverbales Verhalten beschreiben. Sie können so viele Spuren wie nötig für jeden Sprecher anlegen, z.B. drei, wenn zwischen Mimik, Hand- und Körperbewegungen unterschieden werden soll, oder auch gar keine, wenn das nonverbale Verhalten nicht berücksichtigt wird. Es sind auch Spuren dieser Art ohne Sprecherzuordnung (z.B. für Hintergrundgeräusche, Applaus etc.) möglich. Im obigen Beispiel ist die letzte Spur von diesem Typ.
* Typ **'a'** steht für 'annotation' (Annotation). Dies ist der Typ für Spuren, die zusätzliche analytische Information über Ereignisse in den Spuren des Typs **'t'** enthalten. Es kann so viele Spuren dieses Typs wie nötig für jeden Sprecher geben, z.B. eine für die deutsche Übersetzung und eine für die morphologische Transliteration. Im Unterschied zu den anderen Spurtypen existieren Annotationsspuren nicht unabhängig, sondern, da ein Ereignis in einer Annotationsspur, d.h. eine Annotation, immer etwas zu annotieren voraussetzt, von einer Transkritpionsspur abhängig. Ereignisse in diesen Spuren müssen sich daher immer implizit durch zeitliche Zuordnung auf ein Ereignis (oder eine Ereignisabfolge) in der Spur des Typs 't' desselben Sprechers beziehen. Diese Unterscheidung zwischen Transkription und Annotation ist auch ein Teil von dem EXMARaLDa-Datenmodell. In dem obigen Beispiel sind die zweite und die dritte Spur Spuren dieses Typs. Für Ereignisse in diesen Spuren gibt es entsprechende Ereignisse in Spuren des Typs 't', die dem gleichen Sprecher zugeordnet sind: 'Bitte unterbrich mich nicht' stimmt mit den zwei Ereignissen (bzw. der Ereignisabfolge bestehend aus) 'Please do not inter' und 'rupt me' überein und 'loud' mit dem einfachen Ereignis 'you. '. Im Gegensatz dazu, stimmt das Ereignis 'very loud' in der Abbildung unten mit keinem Ereignis in der Transkriptionsspur (Typ 't', Kategorie ‚'v'). Solche Strukturen sind nicht erlaubt und werden zu Strukturfehler führen.



## Segmentketten

Hat man die Definitionen von Spuren, Ereignissen und Spurtypen verstanden, ist die Erklärung von **Segmentketten** unkompliziert: eine Segmentkette ist als eine ununterbrochene Abfolge von Ereignissen in der Spur des Typs 't' definiert. Somit enthält das obige Beispiel zwei Segmentketten - die erste ('Please do not interrupt me. ') besteht aus zwei Ereignissen, die zweite ('I take no orders from you. ') hingegen aus drei. Auf der Grundlage von Segmentketten können Ausgabeformate, die nicht als Partitur dargestellt werden, erstellt werden. Wenn alle Segmentketten anhand ihrer Startpunkte geordnet werden, entsteht eine Drehbuch-ähnliche Darstellung wie die folgende:

**X:** Please do not interrupt me.

**Y:** I take no orders from you.

Segmentketten spielen auch bei der Segmentierung der Basistranskription eine entscheidende Rolle (siehe „How to use segmentation“).

# B. Struktur der Coma-Metadaten

Die Coma-Metadaten bestehen aus fünf vorgegebenen Daten-Containern: Korpora, Kommunikationen, Sprecher, Transkriptionen und Aufnahmen. Zudem können innerhalb dieser Container mehrere Datentypen vorhanden sein. Es ist wichtig, die Beziehungen zwischen diesen Containern und Datentypen zu verstehen.

## Korpora

Korpora sind die Container der höchsten Ebene und beinhalten quasi alle weiteren Container und Datentypen. Sie enthalten auf der nächsten Ebene außer korpusweiten Metadaten und zugehörigen Dateien auch die Container für Sprecher- und Kommunikationen.

## Kommunikationen

Kommunikationen dienen der Beschreibung von Situationen, in denen eine transkribierte Konversation stattfand, sowie der Verwaltung aller zugehörigen Materialien. Meist gehören zu Konversationen Sprecher und es können Aufnahmen und Transkriptionen von der Konversation vorhanden sein. Im Coma-Datenmodell sind Sprecher, Aufnahmen und Transkriptionen den Kommunikationen zugeordnet und mit ihnen verknüpft. Mit der Kommunikation werden außerdem alle relevanten Informationen zur Kommunikation wie Zeit, Ort und Kontext, und gesprochene Sprache(n) gespeichert.

## Sprecher

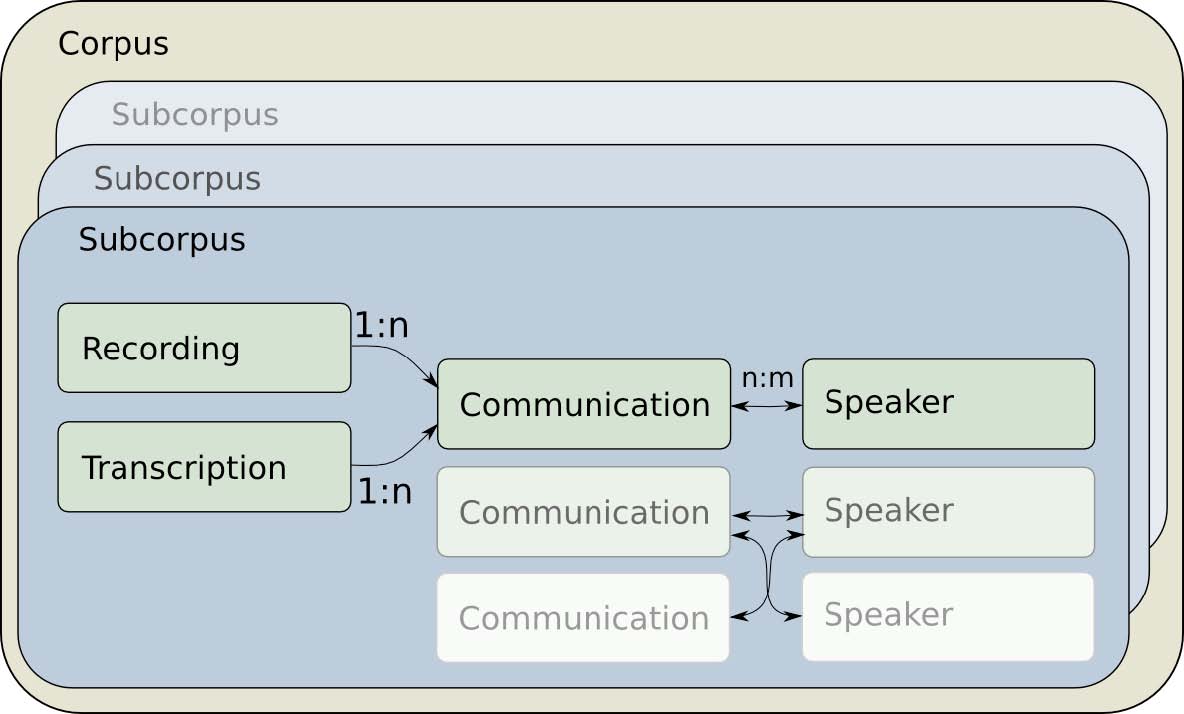
Sprecher sind – wie der Name bereits vermuten lässt – Personen, die an einer Kommunikation teilnehmen. Sprecher müssen keine realen Personen sein (automatisierte Dialogsysteme zählen ebenso) und nicht zwingend sprechen, aber wenn sie für das Verständnis des Kommunikationsgeschehens relevant sind, sollten sie erfasst werden. Der Datentyp Sprecher sollte möglichst viele wichtige Informationen enthalten, z.B. Zeit und Ort der Geburt sowie Sprachlernbiographie. Da Sprecher mehreren Kommunikationen zugeordnet werden können, sollten die für eine bestimmte Kommunikation relevanten Daten nicht mit dem Sprecher, sondern mit der jeweiligen Kommunikation gespeichert werden.

## Aufnahmen

Aufnahmen beziehen sich auf vorhandene Aufnahmen (Audio oder Video) zur Kommunikation. Sie können nicht eigenständig existieren, sondern sind immer mit einer Kommunikation verknüpft.

## Transkriptionen

Transkriptionen stellen die Verknüpfung zu vorhandenen EXMARaLDA-Transkriptionen her, Coma erfasst dabei sowohl Basis-, als auch segmentierte Transkriptionen. In den Voreinstellungen lässt sich allerdings festlegen, ob die Basistranskriptionen angezeigt werden sollen, da z.B. das EXMARaLDA Such- und Analysewerkzeuge „EXAKT“ nur mit segmentierten Transkriptionen arbeiten kann. Genau wie die Aufnahmen müssen auch die Transkriptionen mit einer Kommunikation verknüpft sein. Sie sind über die Kommunikationen mit den transkribierten Aufnahmen verknüpft.



## Weitere Dateitypen

Es existieren weitere Datentypen für die Erfassung von Metadaten. Zwei von ihnen sind besonders wichtig:

### Location

Eine Location repräsentiert einen Ort zu einer bestimmten Zeit.



Eine Location muss nicht zwingend Ort- *und* Zeitangaben enthalten, aber beide sind vorgesehen: im Beispiel gibt die erste Location Auskunft über das Geburtsdatum und den Geburtsort eines Sprechers, die zweite Location hingegen beschreibt nur den Wohnort ohne Zeitangabe. Es ist wichtig, Locations zu verwenden, auch wenn man nur die Zeit(spanne) eines bestimmten Ereignisses erfassen möchte.

### Description

Da es sehr schwierig ist ein vereinheitlichtes Metadatenset für alle Forschungsgebiete zu kreieren, ist die Mehrheit der Metadaten in Coma durch frei wählbare Attribut-Wert-Paare kodiert. Diese Attribut-Wert-Paare befinden sich im Description-Feld, diein allen Coma-Datentypen vorkommen:Korpora, Kommunikationen, Aufnahmen usw. können alle Descriptions enthalten.



Das Beispiel zeigt die Description eines Sprechers. Die Attributnamen innerhalb der Description sind beliebig, dennoch sollten Sie auf ein einheitliches Vokabular in den Metadaten achten. Die Templates in Coma sind hilfreich, um Metadaten konsistent und einfacheingeben zu können.

1. Damit Wortgrenzen als solche erkannt werden, muss hinter jedem Wort ein Leerzeichen gesetzt werden. [↑](#footnote-ref-1)