# Coma Kurzanleitung

Aktuell für Version: 1.5.5



Autor: Kai Wörner (<u>kai.woerner@uni-hamburg.de</u>) Version vom 25. Juni 2009

Die aktuellste Version dieses Dokuments finden Sie unter <a href="http://www.exmaralda.org/files/coma/comadoku.pdf">http://www.exmaralda.org/files/coma/comadoku.pdf</a>

# Über Coma

Der EXMARaLDA Corpus-Manager (CoMa) ist ein Softwarewerkzeug, mit dem EXMARaLDA-Transkriptionen und die dazu gehörigen Aufnahmen mit relevanten Metadaten versehen und zu Korpora zusammengefasst werden. Im Einzelnen beinhaltet der Corpus-Manager folgende Funktionen:

- Eingabe und Verwaltung von Informationen zur Kommunikation, zu den beteiligten Personen, zu den Aufnahmen und zu den Transkriptionen
- Organisation dieser Informationen und Verknüpfung mit den Aufnahme- und Transkriptionsdateien
- Filtern von Metadaten und Zusammenstellung von Teil-/Untersuchungskorpora anhand der Metadaten. Diese Teilkorpora dienen dann als Ausgangspunkt für das Suchwerkzeug EXAKT
- Automatisierte Erstellung von Korpus-Paketen, die sich mithilfe eines Webbrowsers einsehen lassen.

Das Programm befindet sich in einer fortwährenden Weiterentwicklung. Seine Leistungen werden stets ausgebaut.

DIE COMA-DATENSTRUKTUR	4
DIE PROGRAMMOBERFLÄCHE	7
Reiter	7
Menüs	7
Korpus aus Transkriptionen erstellen	9
KORPUSVERWALTUNG	15
VERWALTUNG VON KORPUSMETADATEN	16
Aufbau des Daten-Reiters	16
Verwalten von Kommunikationen	17
Verwalten von Personen und Sprechern	23
Verknüpfen von Kommunikationen und Sprechern	23
Verwalten von Transkriptionen	24
Verwalten von Aufnahmen	26
EINSTELLUNGEN	27

# Die Coma-Datenstruktur

Die Struktur, in der Metadaten in Coma verwaltet werden, kennt einige verschiedene vorgegebene Datencontainer: Korpora, Kommunikationen, Sprecher, Transkriptionen, Aufnahmen. Unter diesen Containern existieren verschiedene Datentypen, die in mehreren dieser Container Verwendung finden können. Da es wichtig ist, das Zusammenspiel dieser Container zu verstehen, sollen sie hier kurz beschrieben werden:

## Korpora:

Korpora bilden die Container für alle anderen Metadatentypen. Korpora enthalten Korpusdaten (also Kommunikationen und Sprecher).

#### Kommunikationen

Mit Kommunikationen werden Gesprächsereignisse verwaltet. An Kommunikationen nehmen typischerweise Sprecher Teil und es kann Aufnahmen und Transkriptionen zu den Kommunikationen geben - im Datenmodell können Aufnahmen, Transkriptionen und Sprecher diesen Kommunikationen zugeordnet sein. Mit den Kommunikationen werden alle Konstellationsdaten zu dem Gesprächsereignis gespeichert, also Dinge wie Ort, Zeit, besondere Umstände, beteiligte Sprachen etc.

## **Sprecher**

Sprecher sind Personen, die sich an der Kommunikation beteiligen. Es muss sich dabei nicht um natürliche Personen handeln (es kann beispielsweise auch ein automatisiertes Dialogsystem als Sprecher geführt werden) und die Sprecher müssen in der Kommunikation auch nicht unbedingt sprechen; wenn sie für die Kommunikation von Belang sind, so werden sie hier erfasst. Die Daten, die zu den Sprechern erfasst werden, sind von der Kommunikation zunächst unabhängig. Es sollten keine Kommunikationsbezogenen Daten mit den Sprechern erfasst werden, damit Sprecher einfach mehreren Kommunikationen zugeordnet werden können. Erfasst werden dagegen Daten wie Geburtsort und -Datum, biographische Angaben, gesprochene Sprachen und deren Status etc.

#### Aufnahmen

Aufnahmen beziehen sich auf Aufnahmen, die zu Kommunikationen angefertigt wurden. Das Coma-Datenmodell unterscheidet zwar Aufnahmen und tatsächliche Mediendateien, diese Unterscheidung ist aber in der Programmoberfläche noch nicht repräsentiert. Aufnahmen lassen sich nur anlegen, wenn es bereits eine Kommunikation gibt, der sie sich zuordnen lassen.

## **Transkriptionen**

In Transkriptionen werden Verknüpfungen zu EXMARaLDA-Transkriptionen gespeichert. Coma verwaltet dabei sowohl Basis- als auch segmentierte Transkriptionen. Zur Anzeige von Basistranskriptionen und zu ihrer Behandlung lassen sich in den #Einstellungen von Coma vorgaben machen. ComaTranskriptionen sind auch - genau wie Aufnahmen - mit Kommunikationen verknüpft - sie lassen sich ohne Kommunikation nicht verwalten. Auch mit den Aufnahmen sind sie nur über die Kommunikation verbunden.



## Weitere Datentypen

Zur Erfassung der eigentlichen Metadaten existieren noch weitere Datentypen, z.B. zur einheitlichen Erfassung von Sprachen und Orten. Zwei dieser Datentypen sind dabei von besonderer Bedeutung:

#### Location

Eine Location bezeichnet in Coma nicht einfach nur einen bestimmten Ort, sondern einen bestimmten Ort zu einer bestimmten Zeit, da biographische und andere relevante Ereignisse häufig einen bestimmten Ort UND eine bestimmte Zeit haben.



In dem oben angegeben Fall wird mit einer Location beispielsweise Geburtsdatum und -Ort eines Sprechers erfasst, in einer weiteren Location nur der Wohnort. Es müssen in Locations nicht immer Ort UND Zeit vermerkt werden, es ist aber wichtig zu wissen, dass man auch eine Location verwenden muss, wenn man nur einen Zeitpunkt vermerken will.

## **Description**

Da es extrem schwierig ist, standardisierte Metadaten für alle erdenklichen Fälle vorzusehen, wird in Coma der überwiegende Teil der Metadaten mit freien Schlüssel-Wert-Paaren erfasst. Diese Paare werden unter dem Datentyp "Description" zusammengefasst. Descriptions existieren für die meisten der in Coma existierenden Datentypen: So lassen sich Descriptions für Korpora, Kommunikationen, Sprecher, Locations, Aufnahmen etc. anlegen.



In diesem Beispiel sind in einer Description Metadaten zu einem Sprecher erfasst. Dass sich die Felder innerhalb von Descriptions frei benennen lassen macht sie nicht beliebig: Jeder Anwender sollte sich für sein Korpus ein festes Inventar an Description-Schlüsseln ausdenken und dokumentieren, um später mit den Metadaten auch vernünftig arbeiten zu können.

# Die Programmoberfläche

## Reiter



Die meisten Funktionen von Coma lassen sich direkt im Programmfenster vornehmen. Das Fenster wird von vier Reitern gegliedert, über die sich die Funktionen für einzelne Bereiche des Programms erreichen lassen.

Mit dem **Korpus**-Reiter kommen sie zur Verwaltung von Korpora und Sub-Korpora. Hier lassen sich die Metadaten, die das Korpus als ganzes betreffen, eingeben und das jeweilige Subkorpus anwählen, das in den anderen Bereichen des Programms bearbeitet werden soll. (Siehe #Korpusverwaltung)

Über den Reiter **Daten** wird der Zentrale Bereich des Programms angewählt. Hier lassen sich die Metadaten zu Gesprächsereignissen, Konstellationen, Sprechern, Transkriptionen und Aufnahmen eingeben, filtern und selektieren. (Siehe #Metadaten)

Der Reiter **Korpus-Korb** bietet Zugriff auf einen "Korb", in dem sich Transkriptionen zur späteren Untersuchung ablegen lassen. Das dahinter liegende Prinzip ist in der Einleitung kurz und im Kapitel Korpus-Korb näher beschrieben.

Unter **Einstellungen** lassen sich verschiedene Programmeinstellungen vornehmen - um welche es sich genau handelt ist unter <u>Einstellungen</u> beschrieben.

## Menüs

In den Programmmenüs finden sich weitere Funktionen, die sich nicht direkt innerhalb des Programmfensters anwählen lassen:

#### Menü Datei



**Neues Dokument:** Erstellt ein leeres Coma-Dokument und legt ein unbenanntes Korpus in ihm an. Nähere Informationen zu Korpora und Subkorpora in Coma finden sie im Kapitel (#Korpusverwaltung)

**Dokument öffnen...**: Zeigt ein Dateiauswahlfenster, über das sich eine Coma-Datei mit Metadaten öffnen lässt.

**Dokument speichern**: Speichert die aktuell geöffnete Coma-Datei. Wenn die Datei noch nicht gespeichert wurde erscheint ein Dateiauswahlfenster, in dem sie den Dateinamen- und Ort angeben können. Beachten Sie dabei die Ausführungen unter (#Korpusverwaltung) für einen sinnvollen Speicherort.

**Dokument speichern als...**: Lässt Sie das Dokument an einem anderen Ort als dem aktuellen Speicherort speichern.

Darauf folgt eine Liste der zuletzt geöffneten Dateien, die sich direkt anwählen lassen.

**Templates öffnen:** Zeigt ein Dateiauswahlfenster, mit dem Sie eine Template-Datei öffnen können. Nähere Informationen zu Templates bekommen Sie bei #Templates.

**Templates speichern:** Speichert die aktuell geöffnete Templates-Datei mit allen angelegten Templates. Sollte keine Templates-Datei geöffnet sein, so erscheint ein Dateiauswahl-Dialog, in dem sich ein Speicherort- und Name für die Datei wählen lässt.

**Templates speichern als...:** Speichert die angelegten Templates in eine Datei, die über einen Dateiauswahl-Dialog bestimmt wird.

**Templates beim Start laden: (...):** Bestimmt, ob die geladene Templates-Datei beim Start von Coma automatisch geöffnet wird.

**Korpus aus Transkriptionen erstellen**: Mit dieser Funktion lässt sich aus EXMARaLDA-Transkriptionen, die bereits mit Kopfdaten versehen sind, eine Coma-Datei erstellen. Diese Funktion ist unter (#Korpus aus Transkriptionen erstellen) beschrieben.

**Programm beenden:** Beendet das Programm.

## Werkzeuge-Menü

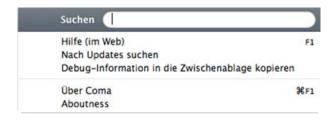


**XML-Dokument anzeigen**: Öffnet ein Fenster mit der XML-Darstellung der Coma-Datei. Beachten Sie, dass es bei grossen Coma-Dateien eine Weile dauern kann, bis das Fenster erscheint.

Mit EXAKT im Korpus suchen...: Öffnet das EXMARaLDA-Suchwerkzeug mit dem aktuell gewählten Korpus. Diese Option funktioniert nur, wenn Sie das vollständige Werkzeugpaket von EXMARaLDA installiert haben.

Nähere Informationen zu EXAKT finden sie unter http://www.exmaralda.org/exakt.html

### Hilfe-Menü



Hilfe (im Web): Öffnet die Coma-Webseite

Nach Updates suchen: Stellt eine Verbindung mit dem EXMARaLDA-Webserver her und überprüft, ob eine neuere Version als die installierte von Coma erschienen ist. Eine derartige Abfrage könnte z.B. so aussehen:



In diesem Fall ist eine neuere Version verfügbar. Mit "Download-Seite besuchen" werden Sie auf die Download-Seite der EXMARaLDA-Homepage geleitet, "Neue Version herunterladen" lädt die aktuellste Version des Programmes herunter, ohne sie zu installieren.

**Debug-Information in die Zwischenablage kopieren:** Kopiert den Inhalt des Log-Files in die Zwischenablage. Wenn Sie einen Fehler im Programm entdecken, so können sie ein Email mit diesem Inhalt an die Programmieren senden.

**Über Coma:** Zeigt ein Fenster mit Programminformationen und der Versionsnummer des Programms an.

## Korpus aus Transkriptionen erstellen

Die Funktion "Korpus aus Transkriptionen erstellen", die sich über das Datei-Menü erreichen lässt, dient dazu, aus vorhandenen EXMARaLDA-Transkriptionen, bei denen bereits Metadaten als Kopfdaten eingegeben wurden, automatisch eine Coma-Datei zu erstellen.

Bevor man diese Funktion benutzt sollte man sich noch einmal klar machen, wie die Dateien eines EXMARaLDA-Korpus sinnvollerweise organisiert sein sollten:

- Mediendateien sollten im selben oder einem untergeordneten Verzeichnis liegen wie die Transkription, der sie zugeordnet sind;
- Transkriptionen, die thematisch oder organisatorisch zusammengehören sollten in Verzeichnissen zusammengefasst sein;
- Die Coma-Datei sollte oberhalb dieser Verzeichnisse gespeichert sein.

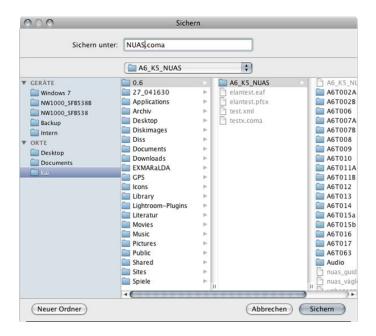
Wenn sie die "Korpus aus Transkriptionen erstellen"-Funktion verwenden, müssen Sie sich um den letzten Punkt nicht kümmern - die Datei wird von dem Assistenten an die richtige Stelle gespeichert.

### Verwenden des Assistenten

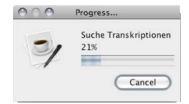
Die Funktion zum Erstellen einer Coma-Datei aus Transkriptionen ist als Assistent aufgebaut, der sie durch die notwendigen Schritte führt. Wenn Sie den Assistenten aufrufen sollten Sie folgenden Bildschirm sehen:



Im Feld "Korpus-Name" geben Sie den Namen ein, den Ihr Korpus später haben soll.



Klicken Sie dann auf die drei Punkte hinter dem Feld "Coma-Datei" und wählen dort **das** Verzeichnis über dem Verzeichnis/den Verzeichnissen aus, in denen Ihre Transkriptionen liegen und wählen Sie einen Dateinamen.



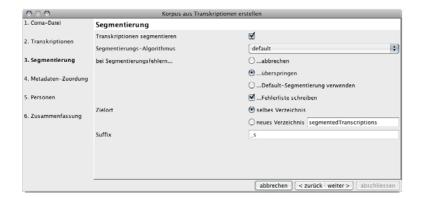
Daraufhin durchsucht der Assistent alle Verzeichnisse unterhalb des gewählten Verzeichnisses nach Transkriptionen und analysiert deren Kopf-/Metadaten. Klicken Sie daraufhin auf die "Weiter"-Schaltfläche, um zum Bildschirm mit der Auswahl der Transkriptionen zu kommen.



In diesem Schritt können Sie festlegen, welche der gefundenen Transkriptionen in die Coma-Datei übernommen werden sollen. In der ersten Spalte der Tabelle können Sie Dateien an- oder abwählen, in der zweiten Spalte ist der Dateiname der Transkription angegeben und in der dritten Spalte wird angezeigt, ob es sich um eine segmentierte Transkription handelt oder nicht. Wollen Sie beispielsweise nur segmentierte Transkriptionen in ihre Coma-Datei übernehmen, so wählen sie zunächst mit der Schaltfläche "alle abwählen" am unteren Rand alle ab und fügen die segmentierten anschliessend mit "Segmentierte Transkriptionen wählen" wieder hinzu.

Beachten Sie, dass die Auswahl der Transkriptionen natürlich einen Einfluss auf die Menge der Metadatenfelder haben kann, die Ihnen zur Verfügung stehen.

Wenn Sie Ihre Auswahl getroffen haben, klicken Sie auf "Weiter".



Im nächsten Bildschirm können Sie wählen, ob Basistranskriptionen automatisch segmentiert werden sollen. Nähere Informationen zu verschiedenen Segmentierungsalgorithmen bekommen Sie in der Dokumentation zum EXMARaLDA-Partitur-Editor.

Das Kontrollfeld Transkriptionen segmentieren bestimmt, ob überhaupt Basistranskriptionen segmentiert werden sollen. Wenn dieses Kontrollfeld abgewählt ist, stehen auch die anderen Optionen nicht zur Verfügung.

Beim Auswahlmenü Segmentierungs-Algorithmus können Sie zwischen den verfügbaren Algorithmen auswählen.

Die nächsten Kontrollfelder bestimmen, was die Segmentierungsroutine machen soll, wenn sich eine Transkription aufgrund eines Fehlers (meist bedingt durch nicht eingehaltene Transkriptionskonventionen) nicht segmentieren lässt:

Bei gewählter Option abbrechen wird der Segmentierungsprozess abgebrochen, sobald sich eine Transkription nicht segmentieren lässt. Auch die vorher bereits segmentierten Transkriptionen werden wieder gelöscht.

Wenn die <mark>überspringen-Option angewählt ist, werden nur die fehlerhaften Transkriptionen nicht segmentiert, die Transkriptionen, bei denen keine Fehler auftreten, werden dagegen geschrieben.</mark>

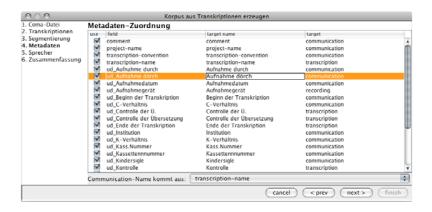
Mit der Option Default-Segmentierung verwenden werden Transkriptionen, bei denen ein Fehler auftritt, mithilfe des Default-Algorithmus behandelt, der zwar keine eigentliche Segmentierung durchführt, aber trotzdem eine Transkription vom Typ "segmentierte Transkription" erstellt.

Mit der Checkbox Fehlerliste schreiben können Sie bestimmen, ob eine Liste mit den Fehlern, die bei der Segmentierung aufgetreten sind, geschrieben wird. Die Datei wird an die selbe Stelle geschrieben wie die Coma-Datei (die Stelle, die am Anfang ausgewählt wurde) und heißt "segmentation-errors.xml".

Unter Zielort können Sie bestimmen, an welche Stelle die segmentierten Transkriptionen geschrieben werden sollen – entweder in das selbe Verzeichnis wie die Basistranskriptionen (selbes Verzeichnis) oder in ein neues Verzeichnis im Stammordner des Korpus. Den Namen des neuen Verzeichnisses müssen Sie in dem Eingabefeld bestimmen.

Mit Suffix bestimmen Sie, was an den Dateinamen der Transkriptionen angehängt wird, um zu kennzeichnen, dass es sich um segmentierte Transkriptionen handelt.

Wenn Sie die Einstellungen vorgenommen haben, klicken Sie auf "Weiter".



Auf dem Bildschirm "Metadaten-Zuordnung" bestimmen Sie, welche der gefundenen Metadaten mit welchem Namen und an welche Stelle in der Coma-Datei übernommen werden.

In der ersten Spalte können Sie bestimmen, ob der jeweilige Schlüssel überhaupt in die Coma-Datei übernommen werden soll oder nicht.

In der zweiten Spalte sehen Sie alle Transkriptions-Kopfdatenschlüssel (außer der Sprechertabellen), die in allen Transkriptionen des Korpus gefunden wurden. Felder, die als "frei bestimmbare Attribute" eingegeben wurden haben in der Liste ein vorangestelltes "ud\_" (für "user defined").

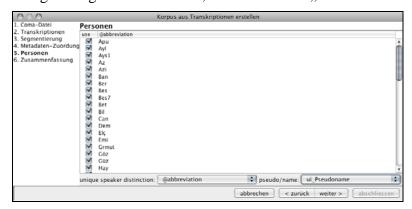
In der folgenden dritten Spalte können Sie den Namen bestimmen, die die entsprechenden Schlüssel in Coma bekommen sollen; in der Regel werden das die vorgegebenen sein, die bereits im Partitur-Editor vergeben wurden. An dieser Stelle lassen sich allerdings auch Daten vereinheitlichen: Im obigen Beispiel ist der Schlüssel "ud\_Aufnahme\_dörch" markiert, der direkt unterhalb des Schlüssels "ud\_Aufnahme\_durch" steht - offensichtlich ein Eingabefehler. In der dritten Spalte liesse sich an dieser Stelle "Aufnahme durch" eintragen, um die Felder in Coma zu vereinheitlichen.

In der vierten Spalte ("Ziel") wird bestimmt, an welcher Stelle in der Coma-Datei die Metadaten schließlich landen: Metadaten, die in einem Transkriptionskopf standen beziehen sich ja nicht notwendigerweise auf eine Coma-Communication. Zur Auswahl stehen "communication" für Metadaten zur Gesprächskonstellation, "recording" für Metadaten zur Aufnahme und "transcription" für Metadaten zur Transkription.

Häufig existieren zu einer Gesprächssituation ("communication") mehrere Transkriptionen. Im Drop-Down-Menü am unteren Bildschirmrand bestimmen Sie, nach welchen Metadatenschlüssel Transkriptionen zu einer Kommunikation zusammengefasst werden - wenn sie also mehrere Transkriptionen in Kommunikationen zusammenfassen wollen, sollten Sie in den Transkriptionen einen Metadatenschlüssel haben, der die entsprechende Kommunikation eindeutig definiert. Wenn jede Transkription genau einer Kommunikation zugeordnet ist, dann wählen Sie "one file -> one communication" aus dem Menü.

Sollten Sie nach dem Schreiben der Coma-Datei feststellen, dass sie an dieser Stelle irgendetwas falsch zugeordnet haben, so können Sie einfach mit dem "zurück"-Knopf wieder an diese Stelle springen und die Zuordnungen ändern - der eigentliche Vorgang des Erstellens der Coma-Datei geht nämlich relativ schnell vonstatten.

Wenn Sie die Zuordnungen vorgenommen haben, klicken Sie auf "weiter".

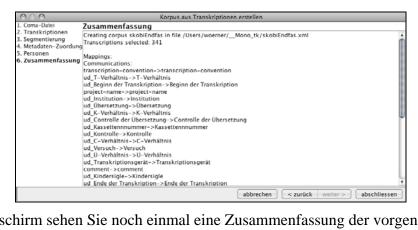


Im nächsten Bildschirm können Sie bestimmen, welche der gefundenen Sprecher in die Coma-Datei übernommen werden soll, indem sie in der ersten Spalte der Tabelle die Sprecher an- oder abwählen. Mit dem linken Drop-Down-Menü bestimmen Sie, anhand welches Metadatenfeldes in den Partitur-Sprechertabellen die jeweiligen Sprecher im Korpus auseinandergehalten werden. In den meisten Fällen wird das Namenskürzel aus der im folgenden abgebildeten Sprechertabelle im Partitur-Editor für die jeweiligen Sprecher eindeutig sein, in den wenigsten Fällen die durchnummerierten Sprecherkürzel ("SPK0, SPK1..."), die im Assistenten als "Id" erscheinen.

Wenn Sie die Auswahl im linken Drop-Down-Menü ändern, so wird die Tabelle mit den zu importierenden Sprechern sofort aktualisiert - so können Sie gleich erkennen, ob Sie das richtige Feld für eine eindeutige Zuordnung ausgewählt haben.

In Coma haben alle Sprecher ein Kürzel und einen (Pseudo-)Namen. Aus welchem Metadatenfeld dieser Name bestimmt werden soll, bestimmen Sie mit dem rechten Drop-Down-Menü.

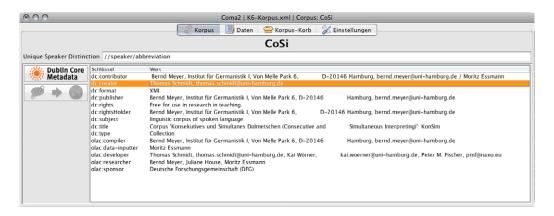
Wenn Sie die entsprechenden Auswahlen vorgenommen haben, klicken Sie auf "Weiter".



Im nächsten Bildschirm sehen Sie noch einmal eine Zusammenfassung der vorgenommenen Einstellungen. Wenn Sie noch etwas ändern wollen, klicken Sie auf die Schaltfläche "zurück" und nehmen Sie Änderungen vor. Ansonsten klicken Sie auf "abschliessen", um die Coma-Datei zu erzeugen. Wenn die Datei fehlerlos geschrieben wurde wird das Fenster geschlossen und die Coma-Datei in Coma geöffnet.

# Korpusverwaltung

Auf dem Korpus-Reiter werden alle Daten verwaltet, die sich mit dem Korpus als ganzes befassen. Der Reiter - hier mit einem geladenen Korpus - sieht so aus:



Zu allen angelegten Korpora lassen sich in Coma beschreibende Metadaten eingeben: Vorgegeben sind dabei der Korpusname und das Feld der "Unique Speaker Distinction" - sie lassen sich im oberen Bereich des Programms editieren.

Das Feld "Unique Speaker Distinction" muss vom Anwender nur geändert werden, wenn die Sprecherabkürzungen für die erfassten Sprecher nicht eindeutig sind und die eindeutige Zuordnung über ein anderes Metadatenfeld erfolgt. In diesem Fall muss der XPath zu diesem Metadatenfeld an dieser Stelle eingegeben werden. Für Korpora, die aus bestehenden Transkriptionen erzeugt werden, wird dieser Wert automatisch generiert.

Zu jedem Korpus lässt sich in der großen Tabelle eine Description (siehe #Datenstruktur) anlegen, in der sich freie Attribut-Wert-Paare zu Eigenschaften des Korpus eingeben lassen. Die Schaltfläche "Dublin Core Metadata" auf der linken Seite der Tabelle fügt die Schlüssel für eine Dublin-Core¹-konforme Ressourcenbeschreibung in die Tabelle ein - eine solche Beschreibung erleichtert die Auffindbarkeit eines Korpus z.B. im Internet. In der Abbildung sind für das ausgewählte Korpus Dublin-Core-Metadaten erfasst.

-

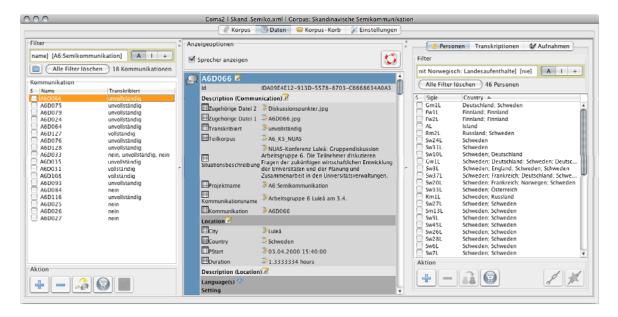
<sup>1</sup> http://dublincore.org/

# Verwaltung von Korpusmetadaten

Alle Metadaten zu Gesprächskonstellationen, Sprechern, Aufnahmen und Transkriptionen werden auf dem Daten-Reiter vorgenommen.

## Aufbau des Daten-Reiters

Der Bildschirm gliedert sich in drei Bereiche: Auf der linken Seite befindet sich eine Tabelle mit den Kommunikationen des gewählten Korpus. Über dieser Tabelle liegt der Bereich, in dem sich Filter für diese Tabelle modifizieren und anlegen lassen (siehe #Filtern). Unter der Tabelle finden sich Schaltflächen zum Hinzufügen, Entfernen und Duplizieren von Kommunikationen sowie für die Bedienung des #Korpus-Korbes und zum #Zählen von Segmenten.



In der mittleren Spalte werden die Metadaten der ausgewählten Kommunikationen oder Sprecher angezeigt und bearbeitet. Über dieser Anzeige findet sich eine Auswahlmöglichkeit ("Sprecher anzeigen"), mit der ausgewählt werden kann, ob in der Anzeige zu ausgewählten Kommunikationen immer auch die zugeordneten Sprecher angezeigt werden sollen.

## Der "Panik-Schalter"

Rechts neben diesem Schalter findet sich eine Schaltfläche mit einem Rettungsring-Symbol. Sollte das Programm einmal nicht reagieren, insbesondere bei der Auswahl von Kommunikationen und Sprechern, so können Sie es mit einem Klick auf diesen Knopf versuchen: Das Programm sollte danach wieder reagieren.<sup>2</sup>

Auf der rechten Seite findet sich eine Spalte, die wiederum drei Reiter hat, mit denen zwischen der Bearbeitung von Personen/Sprechern, Transkriptionen und Aufnahmen umgeschaltet werden kann. Der Inhalt dieser Reiter wird in den entsprechenden Sektionen erklärt.

Die Trenner zwischen den drei Spalten lassen sich verschieben, um mehr Platz für einzelne Spalten auf dem Bildschirm zu bekommen. Am oberen Rand dieser Trenner befinden sich kleine Dreieckemit einem Klick auf diese Pfeile lassen sich Spalten ganz ein- oder ausblenden, so dass sich beispielsweise nur Kommunikationen und deren Metadaten anzeigen lassen.

<sup>&</sup>lt;sup>2</sup> Wir werden natürlich versuchen, den Fehler zu finden, der zu dieser Schaltfläche geführt hat.

## Verwalten von Kommunikationen

Mit den Kommunikationen werden alle Konstellationsdaten zu dem Gesprächsereignis gespeichert, also Dinge wie Ort, Zeit, besondere Umstände, beteiligte Sprachen und Sprecher etc.



Die Kommunikationen werden in einer dreispaltigen Tabelle auf der linken Seite des Reiters angezeigt. In der zweiten Spalte steht der Kommunikationsname, ein in Coma vorgegebenes Feld, in der dritten Spalte erscheint ein wählbares Element aus den Metadaten zu der jeweiligen Kommunikation - im Falle des obigen Screenshots die Information darüber, wie die Kommunikation transkribiert wurde. Die Kontrollkästchen in der ersten Spalte lassen sich vom Benutzer nicht an- oder abwählen; wenn auf der rechten Seite ein oder mehrere Sprecher angewählt werden zeigen diese Kontrollkästchen an, ob diese Sprecher der jeweiligen Kommunikation zugeordnet sind, d.h. an ihr teilgenommen haben.

Mit einem Klick auf die Spaltenüberschriften lässt sich die Tabelle nach dieser Spalte sortieren - dies ist insbesondere für die Warenkorb-Funktionalität hilfreich, aber auch, um sich beispielsweise schnell einen Überblick zu verschaffen, an welchen Kommunikationen ein bestimmter Sprecher teilgenommen hat: Sortieren Sie dazu die Tabelle anhand der ersten Spalte und wählen Sie auf der rechten Seite einen Sprecher aus - die relevanten Kommunikationen stehen jetzt untereinander am Anfang (oder am Ende) der Tabelle.

Mit den Schaltflächen unterhalb der Tabelle lassen sich u.a. Kommunikationen anlegen und löschen.

Ein Klick auf diese Schaltfläche legt eine neue Kommunikation an. Sie erscheint üblicherweise am Ende der Tabelle und hat den Namen "unnamed". Eine eindeutige Id wird automatisch vergeben und kann in Coma nicht geändert werden.

Diese Schaltfläche löscht die ausgewählte(n) Kommunikationen. Achtung: Es erfolgt keine Sicherheitsabfrage und dieser Schritt lässt sich nicht rückgängig machen.

Diese Schaltfläche erstellt eine Kopie der ausgewählten Transkription. Die Kopie erhält eine neue Id und den Namen der Quell-Kommunikation mit einem angehängten "\_cloned", die Metadaten sind ansonsten identisch. Sind beim betätigen der Schaltfläche mehrere Kommunikationen ausgewählt, so wird nur von der obersten eine Kopie erstellt.

Mit dieser Schaltfläche werden alle Transkriptionen, die den augenblicklich ausgewählten Kommunikationen zugeordnet sind, in den "Korpus-Korb" gelegt. Nähere Informationen finden sich im Kapitel zum #Korpus-Korb-Reiter.

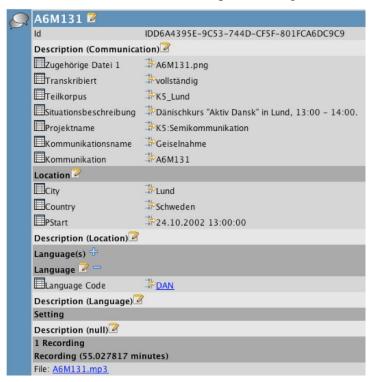
Die Taschenrechner-Schaltfläche führt eine Segment-Zählung durch. Dazu addiert sie die Segmentzahlen aller segmentierten Transkriptionen, die den ausgewählten Kommunikationen zugeordnet sind, und gibt sie in einem kleinen Fenster aus. In diesem Beispiel



haben die (nach HIAT) segmentierten Transkriptionen zu den ausgewählten Kommunikationen beispielsweise 2412 *utterances* und 23210 *words*.

## Metadaten-Anzeige

Wenn eine Kommunikation ausgewählt ist, werden in der mittleren Spalte die erfassten Metadaten angezeigt. Im oberen blauen Balken steht der Name der Kommunikation, der auch immer in der zweiten Spalte der Kommunikations-Tabelle erscheint. Darunter steht die eindeutige Id der Kommunikation, die sich nicht ändern lässt. Es folgen die Description zur Kommunikation und weitere Metadatenfelder, die ebenfalls über eine Description verfügen können.



In der Metadaten-Ansicht befinden sich einige Icons, die sich anklicken lassen:

Mit einem Klick auf dieses Icon wird Wert des Metadatenschlüssels, der sich in dieser Zeile befindet, als zweite Spalte in der Tabelle mit den Kommunikationen angezeigt. Möchte man sich also z.B. einen schnellen Überblick darüber verschaffen, welche Kommunikationen einem bestimmten Projektnamen zugeordnet sind, so würde man neben auf das Symbol neben "Projektname" in der Metadatenansicht klicken.

Mit einem Klick auf das Filter-Icon wird ein Filter für die Kommunikationstabelle erstellt. Klickt man im obigen Beispiel auf das Filter-Icon in der Zeile "Transkribiert - vollständig" klicken, so werden in der Kommunikationstabelle nur noch Kommunikationen angezeigt, bei denen dem

Attribut "Transkribiert" der Wert "vollständig" zugewiesen ist, also nur noch Kommunikationen, die vollständig transkribiert sind. Nähere Informationen zu der Verwendung von Filtern finden sie unter #Filter

Mit dieser Schaltfläche lassen sich die Metadaten editieren. Ein Klick auf das Symbol öffnet je nach Metadatenfeld unterschiedliche Dialoge, in denen sich die entsprechenden Änderungen vornehmen lassen. Die einzelnen Dialoge werden ein wenig später erklärt. #Eingabedialoge

Analog zu den anderen Schaltflächen in Coma lassen sich mit den Plus- und Minus-Schaltflächen Metadatenfelder hinzufügen oder entfernen. Im obigen Beispiel ist ein Plus-Symbol neben der Überschrift "Language(s)"; mit einem Klick auf das Symbol ließe sich der Kommunikation eine weitere Sprache zuordnen.

Mit einem Klick auf das Minus-Symbol neben der Überschrift "Language" liesse sich diese bereits zugeordnete Sprache (Languagecode: "DAN") nach einer Sicherheitsabfrage entfernen.

Es werden keine Metadaten angezeigt, wenn mehrere Kommunikationen ausgewählt sind. In diesem Fall wird nur die Anzahl ausgewählter Kommunikationen und zugeordneter Sprecher angezeigt.

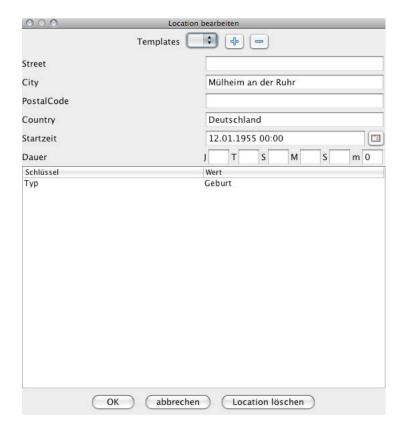
13 Kommunikationen ausgewählt, 140 zugeordnete Sprecher

# **Eingabedialoge**

Wenn in der Metadatenanzeige auf diese Schaltfläche gedrückt wird, erscheint je nach ausgewähltem Metadatenfeld ein Dialog, mit dem dieses Metadatenfeld bearbeitet werden kann. Diese Dialogfelder werden im Folgenden kurz erläutert.

#### Location

Das Location-Dialogfeld dient zur Bearbeitung des Datentyps "Location" – es handelt sich dabei um einen speziellen Datentyp, der auf Seite 5 besprochen ist.



Locations beschreiben – wie dort erwähnt – eine Kombination von Zeit(-abschnitt) und Ort. Die Felder Street, City, PostalCode und Country erklären sich weitgehen von selber. Die (optionale) Startzeit lässt sich bequem durch anklicken des Kalendersymbols an der rechten Seite einstellen: Es erscheint ein Kalender-Auswahlfenster, in dem das Datum nur angeklickt werden muss.



Die Dauer wird in dem Feld eingegeben, das aus mehreren kleinen Eingabefeldern besteht. Diese stehen (von rechts nach links) für: Jahre, Monate, Tage, Stunden, Minuten, Millisekunden. Will man also die Dauer von 2 Jahren und 3 Monaten eintragen, so genügt es, die ersten beiden Felder mit 2 und 3 auszufüllen und dir restlichen Felder leer zu lassen.

Unterhalb findet sich das Feld zur Eingabe der Description, die die Location näher beschreibt; im Falle des Beispiels beschreibt sie nur den Typ des Ereignisses (Geburt).

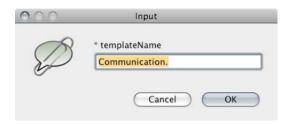
Am oberen Rand des Eingabefensters findet sich ein Panel, mit dem sich Templates anlegen und abrufen lassen. Dies ist auf Seite 22 näher besprochen.

## **Templates**

Für häufig benötigte Datensätze lassen sich in Coma Templates anlegen. Die Datentypen, für die Templates zur Verfügung stehen, erkennt man daran, dass sich ein standardisierter Template-Balken am oberen Rand des Eingabefensters findet.



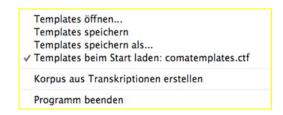
Mit einem Klick auf den Plus-Button werden die Daten, die im Eingabefenster eingegeben wurden, in ein neues Template übernommen. Es erscheint ein Eingabefenster, in dem sich ein Name für das Template eingeben lässt. Danach lassen sich Eingabefenster des selben Datentyps (also z.B. für Communications oder Descriptions) automatisch mit diesen Werten ausfüllen, indem aus dem Dropdown-Menü das Template mit dem vergebenen Namen ausgewählt wird.



Ein Klick auf die Minus-Schaltfläche löscht das im Dropdown-Menü ausgewählte Template.

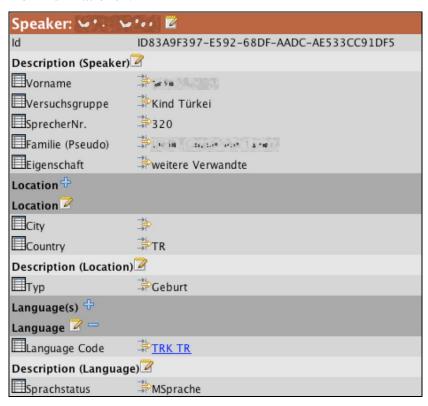
#### **Templates speichern**

Die angelegten Templates lassen sich über das Datei-Menü abspeichern und wieder öffnen. Dort (und in den Einstellungen – siehe auch Seite 27) lässt sich auch festlegen, dass eine bestimmte Template-Datei beim Start des Corpus-Managers automatisch geöffnet wird.



## Verwalten von Personen und Sprechern

Die Verwaltung von Personen und Sprechern erfolgt grundsätzlich analog zu der Verwaltung von Kommunikationen.



Zu beachten ist, dass für Sprecher – im Gegensatz zu Kommunikationen – mehrere Locations angelegt werden können, um Stationen der Biographie von Sprechern abbilden zu können.

## Verknüpfen von Kommunikationen und Sprechern

Um Verknüpfungen zwischen Kommunikationen und den beteiligten Sprechern herzustellen, markieren sie in der Kommunikationen-Tabelle die gewünschten Kommunikationen (auch mehrere - verwenden sie dazu die Betriebssystemspezifischen Tasten zur Auswahl mehrerer Objekte) und in der Sprechertabelle die beteiligten Sprecher und klicken Sie auf die "Verknüpfen"-Schaltfläche (Kettensymbol):



Beachten Sie: Es werden alle selektierten Sprecher zu allen selektierten Kommunikationen hinzugefügt, unabhängig davon, ob sie bereits verknüpft waren oder nicht.

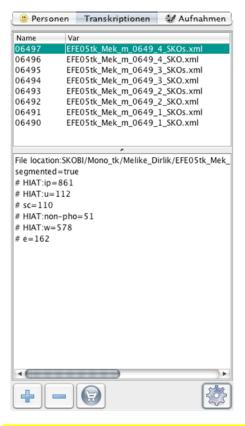
Um Verknüpfungen zwischen Sprechern und Kommunikationen aufzuheben, verfahren Sie analog: Markieren Sie alle Kommunikationen und Sprecher, zwischen denen Sie die Verknüpfung aufheben wollen, und klicken Sie auf die "Verknüpfung aufheben"-Schaltfläche:



Alle Verknüpfungen zwischen den gewählten Kommunikationen werden aufgehoben, ob sie vorher bestanden oder nicht.

ACHTUNG: Beide Aktionen können nicht rückgängig gemacht werden!

## Verwalten von Transkriptionen



Auch Transkriptionen werden in Coma Kommunikationen zugeordnet. Wenn Sie in der linken Spalte eine Kommunikation auswählen, so sehen Sie in der rechten Spalte (sofern Sie den Reiter "Transkriptionen" ausgewählt haben) die Transkriptionen, die dieser Kommunikation zugeordnet sind.

Beachten Sie: In den Einstellungen lässt sich festlegen, ob nicht-segmentierte Basistranskriptionen in Coma angezeigt werden.

Im oberen Teil der Spalte sehen Sie eine Liste mit den zugeordneten Transkriptionen. Wenn Sie eine dieser Transkriptionen auswählen, bekommen Sie im unteren Teil Informationen zu dieser ausgewählten Transkription angezeigt – in gezeigten Beispiel wird der Ort der Transkriptionsdatei im Dateisystem und die Anzahl von HIAT-Segmenten, die in dieser Transkription vorliegen.

Auch an dieser Stelle lassen sich mit dieser Schaltfläche Transkriptionen hinzufügen. Nach einem Klick öffnet sich ein Dateiauswahlfenster, mit dem Sie die Transkriptionen, die zur ausgewählten Kommunikation gehören, auswählen.

Beachten Sie, dass die Transkriptionen im selben Verzeichnis wie die Coma-Datei oder in einem Verzeichnis unter dem Verzeichnis der Coma-Datei liegen müssen.

Wenn Sie die Datei(en) ausgewählt haben, wird die Transkription der ausgewählten Kommunikation zugeordnet. Die Transkription bekommt in Coma einen Namen – dieser entspricht entweder dem Dateinamen oder dem Namen der Kommunikation, der sie zugeordnet ist – welche Option gewählt wird lässt sich in den Einstellungen bestimmen. Bei mehreren Transkriptionen werden an die Namen fortlaufende Nummern angehängt.

Bei der Zuordnung wird in die Transkriptionsdatei hineingeschaut. Dabei wird überprüft, ob es sich überhaupt um eine EXMARaLDA-Transkription handelt, ob es eine Segmentierte- oder eine Basistranskription ist und es werden die Metadaten zu Segmentzahlen ausgelesen. Es findet allerdings keine weitergehende Prüfung auf Korrektheit der Transkription statt.

Die Transkriptionen bekommen auch eine eindeutige ID in Coma, die einer eindeutigen ID in der Transkriptionsdatei entsprechen sollte. Standardmäßig schreibt Coma diese ID bei der Zuordnung in den Kopf der Transkriptionsdatei hinein. Sollten Sie nicht wünschen, dass Coma in die Transkriptionsdateien schreibt, so können Sie das ebenfalls in den Einstellungen abschalten ("Coma-ID beim Zuordnen in Transkriptionen schreiben (sofern nicht vorhanden)").

Mit dieser Schaltfläche können Sie die Zuordnung einer oder mehrerer Transkriptionen löschen. Beachten Sie, dass die eigentliche Transkription nicht gelöscht wird – dies müssen Sie von Hand erledigen.

Mit dieser Schaltfläche befördern Sie die ausgewählten Transkriptionen direkt in den Korpus-Korb. Näheres zum Korpus-Korb finden Sie im entsprechenden Kapitel.

Mit dieser Schaltfläche lassen sich "Aufräumarbeiten" anstoßen, die im Zusammenhang mit verknüpften Transkriptionen stehen. Wenn Sie auf die Schaltfläche klicken, erscheint ein Ausklappmenü, aus dem Sie die gewünschte Aufgabe auswählen können:

Segmentierte Transkriptionen selektieren Transkriptionsdaten neu laden

Die Option "Segmentierte Transkriptionen selektieren" öffnet nacheinander alle dem Korpus zugeordneten Transkriptionen, überprüft, ob es sich um Basis- oder segmentierte Transkriptionen handelt und trägt diese Information in Coma ein. Diese Funktion ist insbesondere für Coma-Dateien interessant, die noch mit älteren Coma-Versionen erstellt wurden und in denen diese Information fehlt.

Die Option "Transkriptionsdaten neu laden" öffnet ebenfalls nacheinander alle Transkriptionen und erledigt dabei die Aufgaben, die auch bei einer manuellen Zuordnung von Transkriptionen ausgeführt werden (siehe oben). Diese Option ist insbesondere interessant, wenn sich Transkriptionen geändert haben und sich diese Änderungen auch in Coma widerspiegeln sollen (z.B. im Bezug auf die Segmentzahlen). Auch lässt sich auf diese Weise herausfinden, ob alle zugeordneten Transkriptionen auch vorhanden sind. Auch bei dieser Option werden IDs in die Köpfe der Transkriptionsdateien geschrieben. Sie können dieses Verhalten in den Einstellungen unter "Coma-IDs bei Wartungsdurchläufen in Transkriptionen schreiben (sofern nicht vorhanden)" abstellen.

Beachten Sie: Bei beiden Optionen erfolgt keine Rückfrage und die gemachten Änderungen lassen sich nicht rückgängig machen (solange Sie die Coma-Datei nicht speichern können Sie aber auch keinen Schaden anrichten. Das Fenster, dass sich bei den beiden Wartungsprozeduren öffnet, kann eine ganze Weile wirken, als würde nichts passieren – haben Sie einfach einen Augenblick Geduld.

## Verwalten von Aufnahmen

Die Verwaltung von Aufnahmen erreichen Sie über den Reiter "Aufnahmen" in der rechten Spalte des Daten-Reiters.



Mit dieser Schaltfläche ordnen Sie eine Aufnahme einer Kommunikation zu. Wenn Sie die Schaltfläche betätigen, erscheint ein Dateiauswahlfenster, in dem Sie die entsprechende(n) Mediendatei(en) auswählen können. Wenn sie die Auswahl bestätigt haben so wird eine Verknüpfung zu dieser Datei innerhalb der Kommunikation angelegt. Die Coma-Datei muss vorher gespeichert sein, da der Link zu der Mediendatei relativ zur gespeicherten Coma-Datei angelegt wird.

Mir dieser Schaltfläche entfernen Sie eine Verknüpfung mit einer Aufnahme. Wählen Sie in der Liste eine oder mehrere Aufnahmen aus und klicken Sie auf die Schaltfläche - die Zuordnung wird darauf ohne Rückfrage gelöscht. Die Schaltfläche ist deaktiviert, wenn keine Aufnahme angewählt ist.

Mit dieser Schaltfläche können Sie eine ausgewählte Mediendatei abspielen, sofern sich auf Ihrem System ein Medienplayer befindet, der das entsprechende Format wiedergeben kann. Coma gibt dabei nur den Befehl zur Wiedergabe der Datei an das Betriebssystem weiter. Die Schaltfläche ist deaktiviert, wenn keine oder mehr als eine Aufnahme angewählt ist.

Mit dieser Schaltfläche aktivieren Sie ein Pop-Up-Menü, mit dem Sie Wartungsarbeiten, die mit Aufnahmen zusammenhängen, aufrufen können. Im Moment existiert nur eine derartige Aufgabe:

Zugeordnete Mediendateien suchen: Mit dieser Option können Sie automatisch alle Mediendateien, die sie in EXMARaLDA-Transkriptionen bereits verknüpft haben, in Coma übernehmen und umgekehrt. Nachdem Sie die Option im Pop-Up-Menü ausgewählt haben erscheint folgendes Fenster:

000	Maintenance	
Zugeordnete	Mediendateien suchen	
☐ In Transkrip	tion(en) übernehmen	
☐ Zuordnung f	ür fehlende löschen	
	abbrechen start	

Wenn sie die Option "In Transkriptionen übernehmen" anwählen, so werden Aufnahmen, die sie innerhalb von Coma mit Kommunikationen verknüpft haben, in den dieser Kommunikation

zugeordneten Transkriptionen ebenfalls verknüpft. Da Transkriptionen nur mit einer Mediendatei verknüpft sein können, wird jeweils nur die erste Mediendatei berücksichtigt.

Wenn sie die Option "Zuordnung für fehlende löschen" anwählen, so werden Zuordnungen in Coma, bei denen sich die Mediendatei nicht finden lässt, gelöscht.

Nach Klick auf den start-Button beginnt Coma, alle Transkriptionen nach verknüpften Mediendateien zu durchsuchen und verknüpft diese dann mir den zugehörigen Kommunikationen innerhalb von Coma.

## Einstellungen

Über den Reiter "Einstellungen" können Sie die Voreinstellungen des Programms ändern. Die einstellbaren Optionen sind im Folgenden erläutert:



## **Sprachdatei**

An dieser Stelle lässt sich die Sprache der Benutzeroberfläche von Coma einstellen.

Die Bezeichnungen der verschiedenen Oberflächenelemente werden in einer Resourcendatei verwaltet - die Deutsche und Englische Version ist weitestgehend vollständig, für die anderen

angebotenen Sprachen sind wir auf die Hilfe der Anwender angewiesen. Wenn sie also gerne bei der lokalisierung von Coma helfen wollen, so melden Sie sich bei

Kai Wörner [kai.woerner@uni-hamburg.de].

Wenn Sie die Sprache der Benutzeroberfläche geändert haben, müssen Sie das Programm neustarten, um die Änderungen zu sehen.

### **Filter**

An dieser Stelle lässt sich einstellen, wie Filter (siehe #Filter) für die Kommunikationen- und Sprechertabellen angelegt werden:

**Filter standardmässig aktivieren:** Mit dieser Einstellung werden Filter sofort angewendet, wenn sie hinzugefügt werden.

Neue Filterergebnisse standardmässig addieren (logisch ,oder' statt ,und'): bestimmt, ob beim Hinzufügen von weiteren Filtern die Ergebnisse logisch ,oder' bzw. ,und' verknüpft werden. Nähere Erläuterungen zu den Auswirkungen finden Sie im Kapitel #Filter.

#### **Benutzerinterface**

**Löschen bestätigen:** Hier können Sie festlegen, ob eine Sicherheitsabfrage erfolgt, wenn Sie Elemente aus den Metadaten löschen wollen.

## Neue Transkriptionen erhalten den Namen der...

Wenn einer Kommunikation eine Transkription zugeordnet wird (siehe #Transkriptionen zuordnen), so erhält diese einen Namen; an dieser Stelle lässt sich auswählen, ob dieser Name dem Dateinamen oder dem Namen der Kommunikation entsprechen soll. Der Dateiname der Transkription wird davon nicht beeinflusst. Wird die Einstellung "Kommunikation" gewählt und eine Kommunikation hat mehrere zugeordnete Transkriptionen, so werden die Namen aufsteigend durchnummeriert.

## **Transkription**

Im EXMARaLDA-System lassen sich die meisten Operationen (Suchen, Visualisierungen etc.) nur mit segmentierten Transkriptionen durchführen. An dieser Stelle können Sie festlegen, ob Basistranskriptionen in Coma überhaupt berücksichtigt werden:

**Basis-Transkriptionen anzeigen:** Wenn Sie diese Option abwählen, so werden in der Transkriptionsansicht nur noch segmentierte Transkriptionen angezeigt.

Nur segmentierte Transkriptionen in den Korb legen: Wenn sie diese Option einschalten, so werden nur segmentierte Transkriptionen in den Korpus-Korb gelegt, unabhängig davon, ob Kommunikationen auch Basistranskriptionen zugeordnet sind.

## Vorlagenverwaltung

Für bestimmte Datentypen in Coma lassen sich Templates (Vorlagen) anlagen – diese sind unter 22 näher beschrieben. An dieser Stelle in den Einstellung können Sie bestimmen, ob eine Template-Datei beim Start von Coma automatisch geladen werden soll und beim verlassen automatisch

wieder gespeichert wird. Mit der Schaltfläche zwischen diesen Optionen können Sie die entsprechende Template-Datei auswählen. Die Option zum automatischen Laden einer Template-Datei findet sich übrigens auch im Datei-Menü des Programms.