

Quickstart – Erstellung von Kollektionen in EXAKT



A. Einleitung

Auf der Basis von Suchanfragen bzw. Suchergebnissen lassen sich im EXMARaLDA Analyse- und Konkordanztool Kollektionen von Transkriptionen erstellen. Hierbei wird für jedes Suchergebnis eine neue Transkription (EXB-Datei) angelegt, für die neue Spuren, die gewünschte Größe (linker und rechter Kontext) und weitere Eigenschaften durch entsprechende Parameter gesteuert werden können.

B. Kollektion erstellen

1. Zur Erstellung einer Kollektion ist zunächst eine Suche in EXAKT auszuführen, die Ergebnisse als Konkordanz liefert. Aktuell können nur Kollektionen aus Suchen in T(ranskriptions)- und A(nnotations)-Spuren erstellt werden.

RegEx (T)	Search:	nicht				
#	S	Communication	Speaker	Left Context	Match	Right Context
1	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	in der zweiten äh Division, sind da auch	nicht	Stammspieler. Wir müssten doch eigentlic
2	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	Mannschaft klar beherrschen. Das ham wir	nicht	!
3	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	ig. (([atmet ein]) Äh • i/ ich kann jetzt	nicht	verstehen, warum die Schärfe reinkommt,
4	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	ch kein Weizenbiertrinker. Ich weiß auch	nicht	, ob wir jetzt • mit dem Stil weitermache
5	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	da, • wo wir waren. Äh warum sollen mer	nicht	zu dem (Punkt) kommen, zu äh/ wo ma scho
6	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	WH	eld gewesen? Und warum haben sie s heute	nicht	gemacht?
7	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	RV	das Allerletzte! • Und ich lass mir das	nicht	mehr so lange gefallen. Das sage ich euc
8	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	RV	abellenführer. Das weißt du, Waldi! Oder	nicht	?
9	<input checked="" type="checkbox"/>	Rudi_Voeller_Wutausbruch	RV	spielt haben! Da konnstest doch früher gar	nicht	hingegangen. Die haben doch Standfußball
10	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	Äh, es geht	nicht	immer äh danach, was der • Arbeitnehmer
11	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	((0,3s)) Nein, äh Sie sollen	nicht	darauf aufpassen, dass die Blumen ((0,3s
12	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	llen aufpassen, ((0,4s)) dass die Blumen	nicht	geklaut werden.((0,5s)) Von äh Eindringl
13	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	ein, nein, nein, das istäh d/ d/ das ist	nicht	erforderlich.
14	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	Ja, warum denn	nicht	?
15	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	Ja nun, Sie können jetzt	nicht	/ äh ich meine also ich habe Ihnen drei S
16	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	((0,4s)) da müssen Sie eine von nehmen,	nicht	? Ich kann (entweder) Sie in den/ (in) di
17	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	((0,5s)) n/ ohne Führerschein geht das ja	nicht	.
18	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	umenwart. (Die anderen) Stellen sind gar	nicht	frei hier in meiner Kartei. ((0,3s))Und
19	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	haben wir auch noch andre Interessenten,	nicht	, die Arbeit (suchen).
20	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	AMT	Da kommen wir gar	nicht	umhin.
21	<input checked="" type="checkbox"/>	Helge_Schneider_Arbeitsamt	KLA	Da muss ich aufpassen, dass die Blumen	nicht	abhauen, oder wat?

2. Erstellung der Kollektion starten über **Concordance > Create collection from concordance...**
3. Im erscheinenden Dialog können Parameter gesetzt werden, von denen einige obligatorisch sind:

Create collection

Left context (in sec.)

Right context (in sec.)

(Optional) Annotate matches with (use # for numbering)

Reset start time to 0.0

☐

(Optional) Name of collection (used for result files)

Output directory

Select ...

(Optional) Template file

Select ...

OK

Cancel

Left context (in sec.)	Inhalte innerhalb des angegebenen Zeitraums in Sekunden VOR dem Suchergebnis (Match), werden in die neue Transkription übertragen, z.B.: 15
Right context (in sec.)	Inhalte innerhalb des angegebenen Zeitraums in Sekunden NACH dem Suchergebnis (Match), werden in die neue Transkription übertragen, z.B.: 15
(Optional) Annotate matches with (use # for numbering)	Falls ein Annotationstext angegeben wird, wird in den neuen Transkriptionen eine Annotationsspur angelegt, in der die Matches aus dem Suchergebnis annotiert wird. Der Platzhalter # fügt die Position des Suchergebnisses in der Konkordanz ein, z.B. <i>MATCH #</i>
Reset start time to 0.0	Bei Setzen eines Häkchens, wird die Startzeit der neuen Transkription auf 0 gesetzt und alle nachfolgenden Zeitpunkte entsprechend neu kalkuliert. Ansonsten werden die Zeiten aus der Original-Transkription übernommen.
(Optional) Name of collection (used for result files)	Hier kann ein Name für die zu erstellende Kollektion angegeben werden. Dieser wird für die Generierung der Dateinamen der neuen Transkripte verwendet, z.B.: <i>meineKollektion - Corpus-meineKollektion-1.exb</i>
Output directory	Auswahl eines (neuen) Verzeichnisses, in das die neuen Transkriptionen geschrieben werden. ACHTUNG: Vorhandene Dateien werden überschrieben!
(Optional) Template file	Optionale Auswahl einer Template-Datei mit Definitionen in die neue Transkription einzufügende zusätzliche Spuren (s.u.)

4. Nach Setzen der Parameter kann der Dialog mit „Ok“ bestätigt werden. Die Erstellung der Kollektion erfolgt im angegebenen Ausgabeverzeichnis und kann in Abhängigkeit der Trefferanzahl in der Konkordanzsuche einige Zeit in Anspruch nehmen.

C. Template-Datei

Den Transkriptionen in der Kollektion können automatisch vordefinierte Spuren angefügt werden, indem in der Auswahl „(Optional) Template file“ eine zuvor angelegte XML-Datei (z.B. benannt als tier-template.xml) mit folgender Struktur ausgewählt wird:

```
<?xml version="1.0" encoding="UTF-8"?>
<tier-template>
  <tier type="a" category="anno-new"/>
  <tier type="d" category="desc-new"/>
</tier-template>
```

Für jede anzulegende Spur wird ein Element <tier .../> angegeben, welches im Attribut type den Spurtyp (z.B. „a“ für Annotation, „d“ für Description) und im Attribut category den Namen der Spur enthält.