

Z2-Tagger für EXMARaLDA-Transkriptionen

Der Z2-Tagger für EXMARaLDA-Transkriptionen ist eine Modifikation des Z2-Taggers für TEI-Texte (<http://www.exmaralda.org/z2tagger>) und dient als *proof-of-concept*-Implementation für das System der EXMARaLDA-Standoff-Annotationen (Sextant, <http://www.exmaralda.org/sextant>).

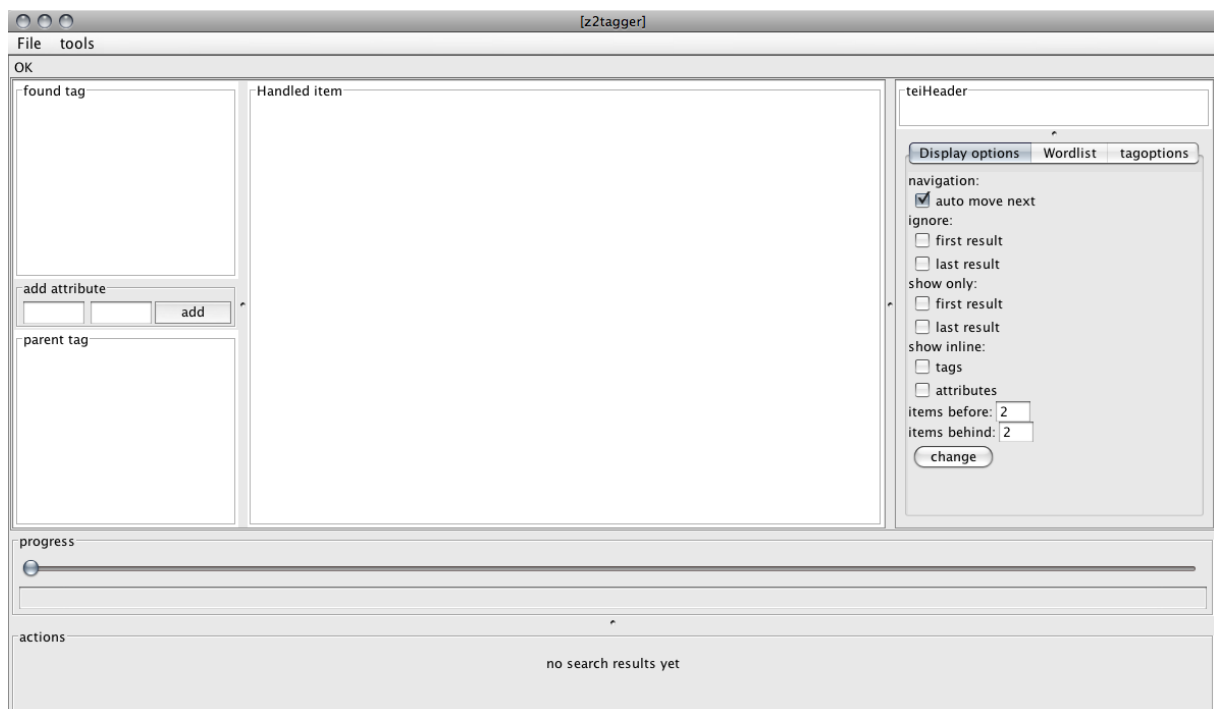
Dieses Dokument dient als kurze Einführung in das Programm – es ist kein ausführliches Manual.

Starten des Programms

Um das Programm zu starten benötigen Sie eine funktionierende Java Laufzeitumgebung der Version 6 oder höher. Wie Sie diese bekommen können erfahren Sie z.B. unter <http://www.exmaralda.org/faq.html>.

Das Programm kann als ausführbares Java-Paket unter <http://www.exmaralda.org/sextant> heruntergeladen werden. Wenn Sie Java richtig installiert haben, startet ein Doppelklick auf das Programmsymbol die Anwendung.

Sie sollten ungefähr den folgenden Bildschirm sehen:



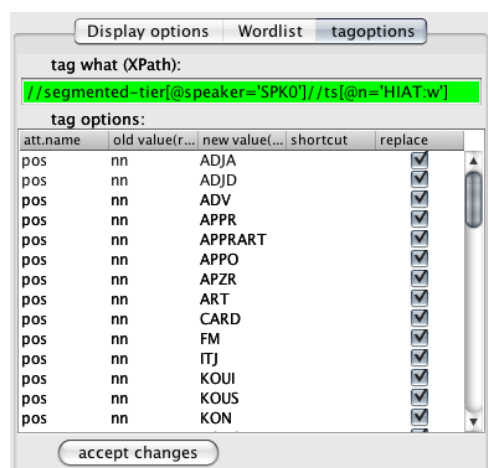
Das Programm benutzen

Um das Programm zu benutzen, öffnen Sie zunächst eine segmentierte EXMARaLDA-Transkription über das File-Menü:

open EXS-file open EXS-file: /Users/kai/Diss/Sextant/wehtun_s.exs open tagging options open tagging options: /Users/kai/Diss/Sextant/taggeroptions.xml
attach stylesheet
save TEI-File save TEI-File as...
save tagging-options save tagging-options as...
save exmaralda standoff...

Wenn die Transkription erfolgreich geladen werden konnte, erscheint der Name der Datei in der Titelleiste des Programmfensters.

Legen Sie nun die Optionen fest, mit denen Sie bestehenden Segmenten der Transkription Annotationen hinzufügen möchten. Legen Sie dazu zunächst die Segmenttypen fest, die Sie annotieren wollen.



Dazu müssen Sie auf der rechten Seite den Reiter „tagoptions“ wählen, und den entsprechenden Xpath zu diesem Segment in das Eingabefeld „tag what (XPath)“ eingeben. Im gezeigten Beispiel zeigt der Pfad

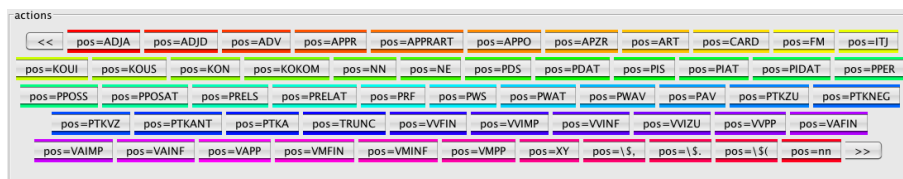
```
//segmented-tier[@speaker='SPK0']/ts[@n='HIAT:w']
```

auf alle HIAT:w (=Wörter der HIAT-Segmentierung)-Segmente des Sprechers *SPK0*.

Wenn Sie auf die Schaltfläche „accept changes“ klicken, färbt sich das Eingabefeld entweder grün (der XPath ist korrekt) oder rot (der XPath ist fehlerhaft).

In der Tabelle darunter („tag options“) legen Sie die Attribut-Wert-Paare fest, die Sie als Annotationen verwenden wollen. In der Spalte „att.name“ legen Sie den Attributnamen (hier pos für part of speech) fest. In der dritten Spalte legen Sie den Wert fest, den das Attribut bekommen soll. In der letzten Spalte legen Sie fest, ob bestehende Werte überschrieben werden sollen. Für alle möglichen Werte eines Attributes legen Sie jeweils eine derartige Tabellenzeile an.

Wenn Sie alle Optionen eingetragen haben, klicken Sie wieder auf „accept changes“. Daraufhin wird für jede Zeile, die sie in der „tag options“-Tabelle eingetragen haben, eine Schaltfläche am unteren Rand des Fensters unter „actions“ erstellt, mit der Sie später das Attribut-Wert-Paar als Annotation anwenden können:

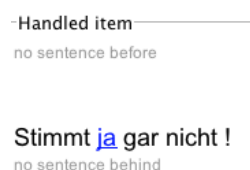


Sie können die gesetzten Parameter und Einstellungen über das Dateimenü mit dem Punkt „save tagging options (as)“ abspeichern, um sie später für andere Transkriptionen wiederverwenden zu können.

Wenn Sie eine Transkription geladen, alle Einstellungen vorgenommen und auf „accept changes“ geklickt haben, wird in der Statuszeile am oberen Bildschirmrand angezeigt, wieviele Segmente auf den Xpath-Ausdruck passen:



Sie können sich jetzt mit den Navigationsschaltflächen, die sich am Anfang und am Ende der Schaltflächen für die Annotationsoptionen befinden, durch die Suchergebnisse bewegen:



Das jeweils aktuelle Wort wird unterstrichen dargestellt. Sollte es bereits mit einer Annotation versehen worden sein, so wird es farbig dargestellt. Mit einem Klick auf eine Schaltfläche am unteren Rand fügen Sie dem Suchergebnis eine Annotation hinzu. Sollte die Option „auto move next“ eingeschaltet sein (sie findet sich unter dem Reiter „Display options“), so springen Sie anschließend direkt zum nächsten Suchergebnis.

Um die Annotationen in einer EXMARALDA-Standoff-Annotationsdatei abzuspeichern, wählen Sie „save exmaralda standoff“ aus dem File-Menü und speichern Sie die Datei mit der Dateierdung „esa“ in dasselbe Verzeichnis wie die Transkriptionsdatei.

Diese Annotationsdatei lässt sich nun mit dem Sextant-Werkzeug validieren.

Nähere Informationen zu weiteren Optionen des Programms (die sich aber vorwiegend auf die Annotation von TEI-Texten beziehen) finden sich unter <http://www.exmaralda.org/z2tagger>.