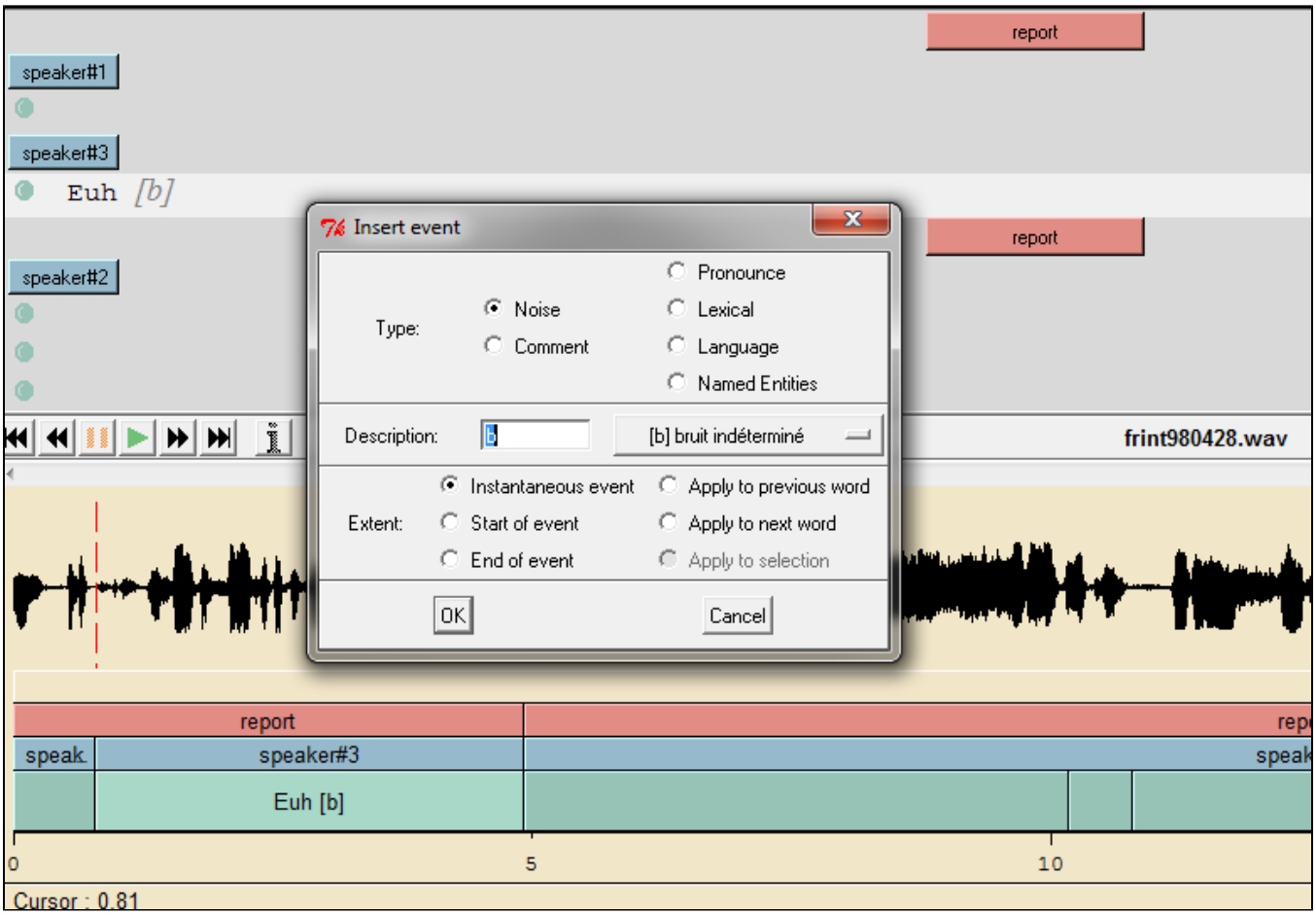
Tools und Manuals im TdR-Wiki

Tools

(Text-)Daten bringen keinen Nutzen, wenn nicht in irgendeiner Weise mit ihnen gearbeitet werden kann. Hierzu lässt sich eine Vielzahl von Tools benutzen, von denen einige hier Erwähnung finden sollen.

Einige davon dienen der Transkription und der anschließenden Annotation (Praat, Transcriber) von Daten. Eine Audiospur kann mit ihrer Hilfe segmentiert und mit Text oder zusätzlichen Informationen versehen werden:



Eine Kurzanleitung zur Benutzung des Transkriptionsprogramms Transcriber ist in das im TdR-Wiki integriert.

Die Möglichkeit einer Wortartenannotation (POS-Tagging) bieten Programme wie TreeTagger, Morphalou (Französisch) und MORPH-IT (Italienisch), welche Wörter erkennen und ihnen sowohl die Wortartinformation als auch das Lemma beifügen.

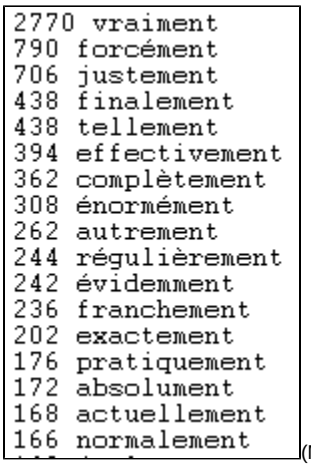
##Screenshot Morph-IT###

Eine weitere Kategorie vereint einige Konkordanzprogramme (AntConc, SCP, TextSTAT), welche in der Lage sind, sprachliche Kontexte von Suchbegriffen auszugeben.

```
  382          même/je/l'/aime/et/eh/parce/que/j'/ai/eu/l'/occasion
  732        travailler/à/Orléans/euh/parce/que/Orléans/m'/attirait/
  789        un/petit/peu/spécial/pas/parce/que/je/vais/vous/expliquer
  945      moyenne/pour/oui/cette/raison/parce/que/je/gère/deux/supermarchés
 1184      arrivent/sept/heures/du/matin/parce/qu'/on/se/lève/très/tôt/
 1275          boucheries/en/supermarché/parce/que/j'/ai/donné/cet/esprit
 2089       métier/j'/ai/pas/pu/le/faire/parce/que/malheureusement/étant
 2147      vas/t'/en/aller/à/tel/endroit/parce/qu'/il/y/a/une/place/de/
 2291           -là/euh/c'/est/peut-être/parce/que/c'/est/mon/tempérament
 2344     enfin/je/suis/plutôt/contre/parce/que/euh/j'/ai/eu/l'/expérience
 2355     eu/l'/expérience/à/la/maison/parce/que/euh/ma/femme/était/avec
 2392        à/la/maison/je/le/faisais/parce/que/j'/estime/que/quand/
 2576      bah/le/samedi/oh/pas/question/parce/qu'/on/se/lève/à/quatre/
 2677       mon/dieu/on/est/revenu/quoi/parce/qu'/il/faisait/mauvais/d'
 2799       rousse/hein/faut/pas/croire/parce/que/je/dis/ça/que/tous/les
 2862       pas/oui/prendre/de/vacances/parce/que/on/a/un/métier/qui/se
 2917        ça/me/ça/ça/me/chagrine/ça/parce/que/ça/fait/huit/ans/que
 3013          qui/travaillait/le/cheval/parce/que/j'/ai/pas/trouvé/de/
```

(Konkordanzen von *parce que* in .txt-Format, ausgegeben vom Simple Concordance Program (SCP))

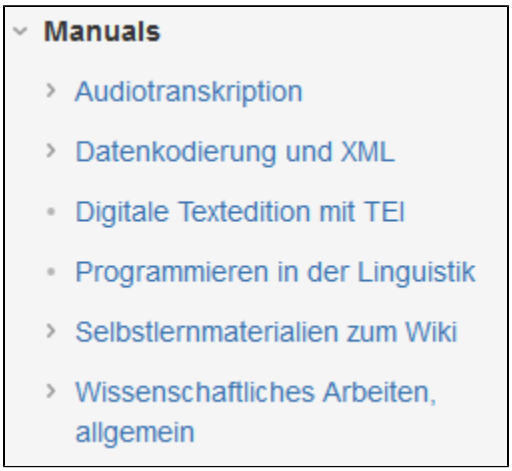
Auch lassen sich beispielsweise mit dem SCP Frequenzlisten erstellen:



(Mit SCP erstelle Frequenzliste von Adverbien auf *-ment*)

Will man einen Eindruck von bestimmten sprachlichen Erscheinungen statistisch verifizieren, so helfen diese Tools bei der Quantifizierung von normalerweise qualitativen Daten.

Manuals



Für einige Tools der Audiotranskription (Transcriber, Praat in Arbeit) stellt das TdR-Wiki Kurzanleitungen zur Verfügung. Diese vermitteln grundlegende Kenntnisse im Umgang mit entsprechenden Programmen, bezwecken jedoch keine vollständige Darstellung der gegebenen Toolfunktionen. Andere Manuals vertiefen die Arbeit an Textdaten, behandeln zum Beispiel die Kodierung von Texten nach dem TEI-Standard oder mithilfe der XML-Auszeichnungssprache. Dieser Bereich eignet sich insbesondere für fortgeschrittenere Studierende, welche hier lernen, wie man Textdaten kodiert, manipuliert und analysiert. Ebenfalls an Fortgeschrittene richtet sich das Manual *Programmieren in der Linguistik*, welches in den Umgang mit der Programmiersprache Python in einem linguistischen Kontext einführt. Ein weiterer Bereich ist dem wissenschaftlichen Arbeiten im Allgemeinen gewidmet und stellt Informationen zu wissenschaftlichen Konventionen zusammen. Auch das vorliegende Manual zur Wikidarstellung ist hier zu finden.

Unbekanntes Makro: 'scrollbar'