



SpeechRecorder – skriptgestützte Sprachaufnahmen im Feld und im Studio

Christoph Draxler
draxler@phonetik.uni-muenchen.de

31.03.2014

Sprachdatenbank

Eine Sprachdatenbank besteht aus

Primärdaten Audio-, Video- und Sensordaten gesprochener Sprache

Sekundärdaten Annotationen, abgeleitete Signaldaten

Tertiärdaten Metadaten, Protokolle, Berichte, Verträgen, usw.

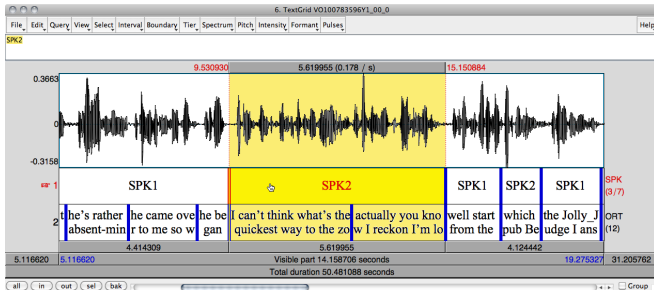
Sie hat eine sichtbare Struktur, erlaubt Browsen, Suchen sowie Datenanalysen und ist (hoffentlich) langfristig und online verfügbar (Draxler 2008, S. 13).

Mindestanforderungen

Sprachsignal	qualitativ optimale digitale Aufnahmen	unkomprimiert, WAV
Transliteration	orthografisch, Chunk-aliniert, theorieneutral	UTF-8 Text
Lexikon	Aussprache aller Tokens der Transliteration	IPA, SAM-PA
Metadaten	Beschreibung der Aufnahmen, Sitzungen und Sprecher	CMDI

Chunk-aliniert

Zeitgrenzen des (Roh-)Transkripts im Segment markieren

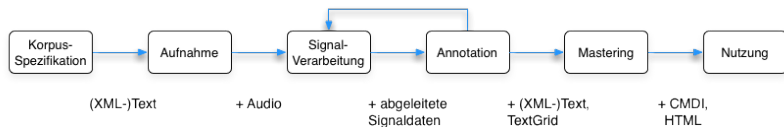


- erlaubt rasches Auffinden der ungefähren Stelle
- macht automatisches Segmentieren langer Signale erst möglich

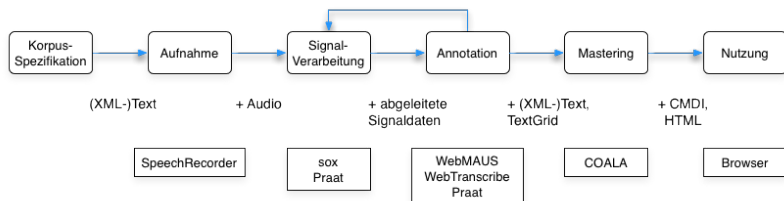
Workflow



Workflow und Formate



Workflow, Formate und Tools



Siehe auch

- www.bas.uni-muenchen.de/forschung/BITS/TP1/Cookbook/
- www.bas.uni-muenchen.de/forschung/BITS/TP2/Cookbook/
- (Draxler 2008) mit Beispiel Ph@ttSessionz in Kap. 9 und 10

Rechtliche Aspekte

WER darf mit WELCHEN Daten WAS machen?

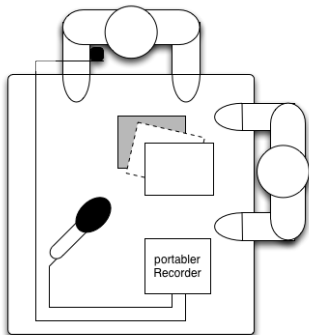
- Widersprüchliche Anforderungen
 - anonyme Sprachaufnahmen
 - Einverständnis der Sprecher muss überprüfbar sein
 - Abrechnung erfordert namentlich bekannte Sprecher
- BAS und CLARIN bieten Muster-Erklärungen
- ggf. muss die Ethik-Kommission zustimmen

www.phonetik.uni-muenchen.de/Bas/BasTemplateInformedConsent_de.pdf

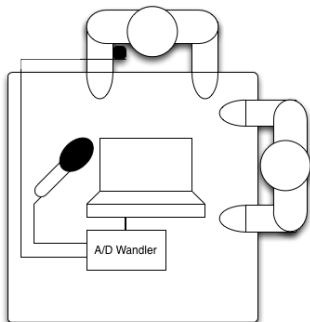
Sprachaufnahmen

- sind die entscheidende Phase
 - alle weiteren Phasen hängen von den Aufnahmen ab
 - Fehler können später kaum mehr korrigiert werden
- sind handwerkliche Arbeit
 - man kann sie lernen
 - Erfahrung und Routine sind wichtig
- sind eine soziale Aktivität

Aufnahmesituationen einfach (schematisch)



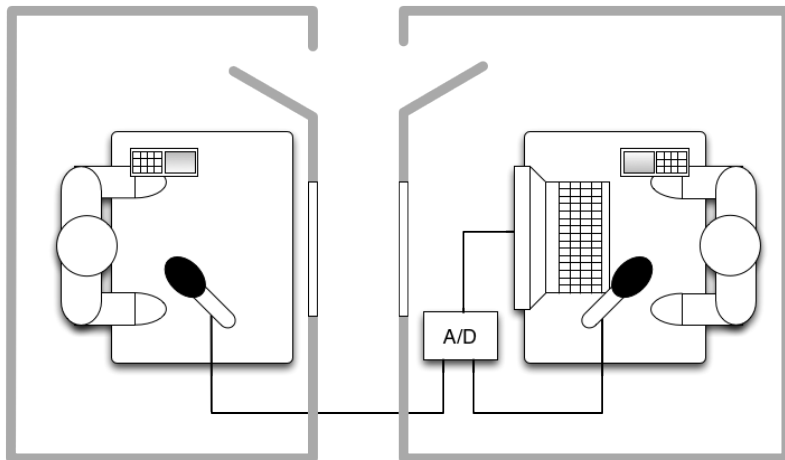
a)



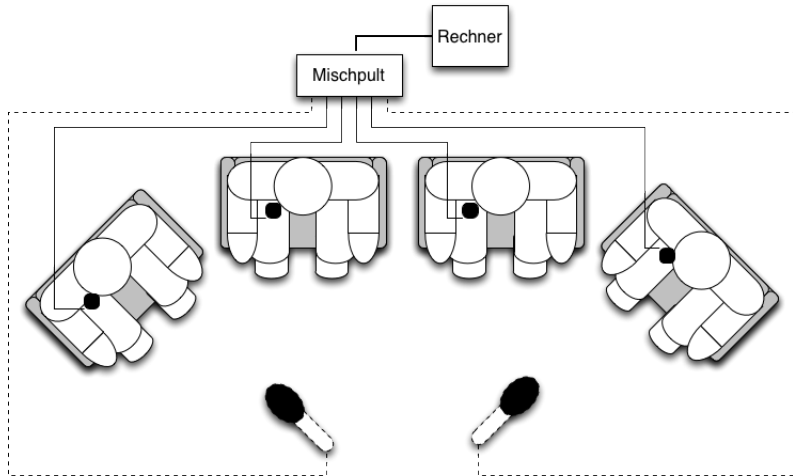
b)

a) mit portabilem Recorder, b) mit Rechner

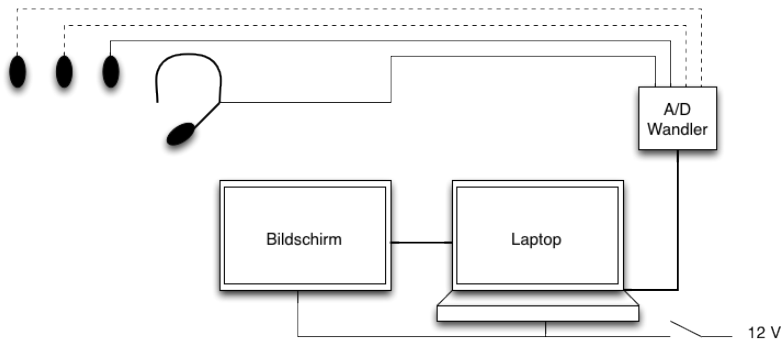
Aufnahmesituation: zwei Räume



Aufnahmesituation: Meeting, Dialog



Aufnahmesituation: mobile Umgebung

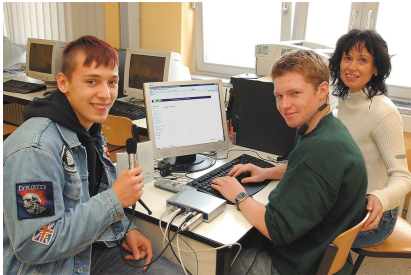


Aufnahmesituationen (echt)



'Deutsch heute' Aufnahmen des IDS mit Stefan Kleiner

Aufnahmesituationen (echt)



Ph@ttSessionz Aufnahmen in Schwerin und Nürnberg



Aufnahmesituationen (echt)



Learners Corpus Aufnahme im Studio des IPS München

Signalqualität

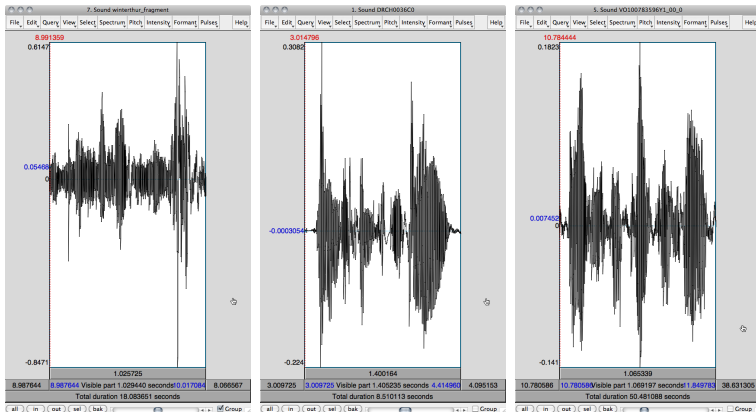
Optimale Signalqualität ist wichtig!

- digitale Geräte verwenden
- Samplerate 48 kHz, 16 bit linear, WAV-Format
- Mikrofon entsprechend der Situation, möglichst nah am Sprecher

Signalqualität lässt sich nachträglich nicht verbessern

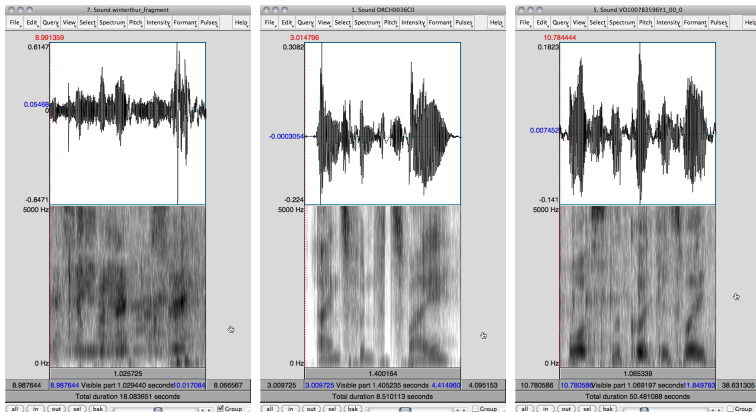
- Glauben Sie an Wunder, aber bitte nicht hier!

Signalqualität



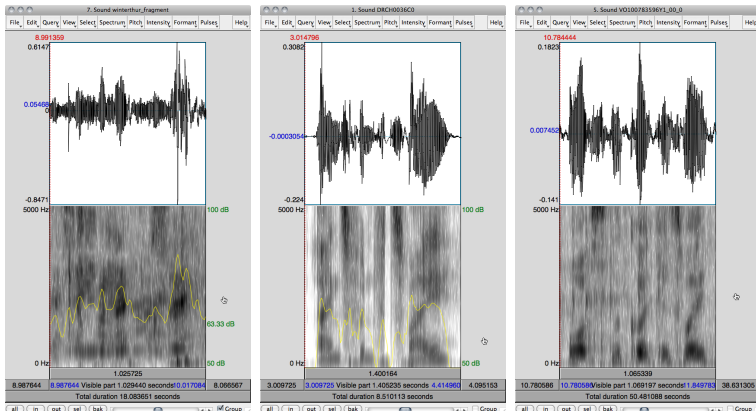
kaum ein Unterschied zwischen den Kurven: zur Beurteilung der Qualität reicht das Oszillogramm nicht aus...

Signalqualität



unterschiedlich starke Kontraste bzw. Rauschen: das Sonagramm hilft etwas...

Signalqualität



Intensität – die gelbe Kurve über dem Sonogramm – liefert weitere Hinweise...auf dem Schulhof ist es laut!

Erfahrungen

schlechtes Signal = aufwendige Transkription

- 1 Minute gelesener Text im Studio → 2-5 Minuten Transkription
- 1 Minute Freundinnen-Dialog auf dem Schulhof → ca. 25 Minuten Transkription

motivierte Sprecher sind geduldig

- immer auch zusätzliches standardisiertes Material aufnehmen (Ziffern, Zahlen, Datum, Geldbeträge, Ortsnamen)

Hardware ist billig

- mindestens zwei Backups (Speicherkarten und ext. Disk)
- lieber viele kleine (1-2 GB) Speicherkarten als eine große



SpeechRecorder Übersicht

- Plattform-unabhängig: Windows, Mac OS X, Linux
- Text-, Bild- und Audio-Prompts
- Sprecher- und Versuchsleiteransicht
- flexible Aufnahmeskripte
- präzise Aufnahmesteuerung
- Online-Aufnahmen mit WikiSpeech
- Handbuch

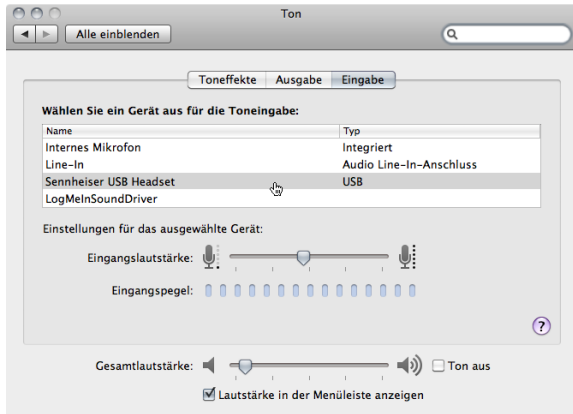
<http://www.phonetik.uni-muenchen.de/Bas/software/>

Blitzstart in 5 Schritten

- 1 SpeechRecorder herunterladen und installieren
- 2 Projekt -> Öffnen und darin 'CLARIN_Workshop' anklicken (oder von außen importieren)
- 3 Sprecher auswählen oder neuen Sprecher hinzufügen
- 4 Projekt -> Einstellungen öffnen und Audio-Einstellungen vornehmen
- 5 Button Aufnahme anklicken und los geht's

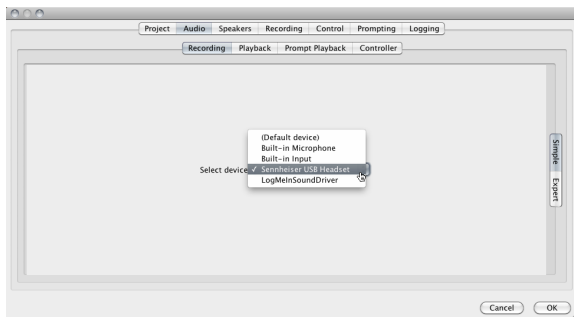
Die Aufnahmen liegen im Verzeichnis RECS im Projekt-Verzeichnis

Audio-Einstellungen im System



hier: Mac mit Sennheiser Headset. Alternative...

Audio-Einstellungen in SpeechRecorder

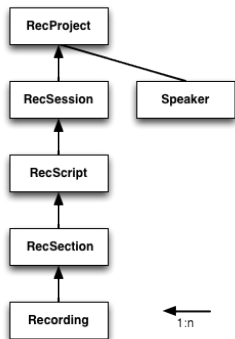


im Menü Projekt -> Einstellungen -> Audio -> Recording

Aufnahmeskript

- definiert den Inhalt einer Aufnahmesitzung
- ist ein wohlgeformtes XML-Dokument
- erstellen und anpassen mit
 - SpeechRecorder, jEdit, Emacs, TextWrangler o.ä.
 - Word und Publishing-Software sind *nicht* geeignet!
- man kann mehr als ein Skript pro Projekt haben
 - wählen Sie das passende Skript im Menü Projekt -> Einstellungen

Skript Details



RecProject eine oder mehrere Sitzungen und Sprecher

Speaker Sprecherdaten

RecSession ein oder mehrere Skripten

RecScript unterteilt in sequentielle Abschnitte

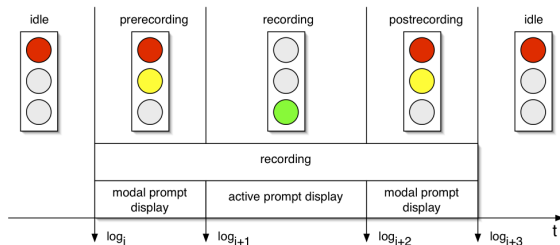
RecSection Aufnahme-Items sequentiell oder randomisiert

Recording Aufnahmeeinstellungen (Code, Steuerung, etc.)

Aufnahme-Item in XML

```
<recording itemcode="L1" postrecdelay="500"  
  prerecdelay="500" recduration="20000">  
  <recinstructions mimetype="text/plain">  
    Bitte buchstabiere...  
  </recinstructions>  
  <recprompt>  
    <mediaitem>  
      M Ü N C H E N  
    </mediaitem>  
  </recprompt>  
</recording>
```

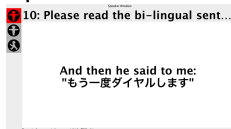
Aufnahmephasen



idle	keine Aufnahme
pre-, postrecording	Umgebungsgeräusch aufnehmen
recording	Sprachaufnahme

Versuchsleiter-Ansicht

- Ansicht des Sprecherfensters



- Signal-Display



Mikrofonpegel mit Spitzen-Anzeige

- Details

ID: 00... Co... ABC Na... Test Vo... Test
Ge... F Ak... FR Ge...

Recording Progress Status

L...	File	Prompt	Re...
0	demo_0...	Welcome to the Speech...	<input type="checkbox"/>
1	demo_0...	and here is the prompt...	<input type="checkbox"/>
2	demo_0...	The recording script is...	<input type="checkbox"/>
3	demo_0...	In the next section, a s...	<input type="checkbox"/>
4	demo_0...	M O R G O N S T U N D	<input type="checkbox"/>
5	demo_0...	How did you get here L...	<input type="checkbox"/>
6	demo_0...	2 7 4 1 6 8 3 9 5 0	<input type="checkbox"/>
7	demo_0...	A Paris il y a 14 lignes...	<input type="checkbox"/>
8	demo_0...	Det är vackert värde...	<input type="checkbox"/>
9	demo_0...	Morgenstund hat ...	<input checked="" type="checkbox"/>
10	demo_0...	And then he said to me ...	<input checked="" type="checkbox"/>
11	demo_0...	Maanantai 24. huhtiku...	<input type="checkbox"/>
12	demo_0...	Mits kult mieliten sa	<input type="checkbox"/>

Aufnahme-Items sind als aufgenommen markiert, können aber überschrieben werden

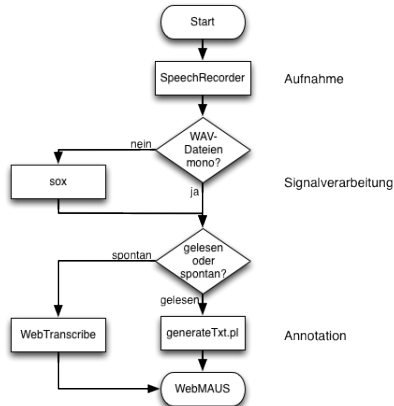
- Aufnahmesteuerung



Arbeitsschritte im Workshop

- ① Sprachaufnahme mit SpeechRecorder
- ② prüfen, ob alle WAV-Dateien nur einen Kanal haben
- ③ für jede WAV-Datei die passende TXT-Datei anlegen
- ④ die Aufnahmeverzeichnisse mit WebMAUS automatisch segmentieren
- ⑤ Metadaten aus den Aufnahmen und -verzeichnissen erstellen
- ⑥ CMDI-Dateien in Repository ablegen

Arbeitsschritte im Workshop



Verwandte Software



WikiSpeech Audioaufnahmen über das Web



WebMAUS Automatische Segmentation



Percy Online-Perzeptionsexperimente

<http://www.phonetik.uni-muenchen.de/Bas/software/>

Erstellung und Validierung von Sprachdatenbanken

- www.bas.uni-muenchen.de/forschung/BITS/TP1/Cookbook/
- www.bas.uni-muenchen.de/forschung/BITS/TP2/Cookbook/
- Chr. Draxler (2008) Korpusbasierte Sprachverarbeitung. Narr Studienbücher, Narr Verlag, Tübingen

SpeechRecorder

- Chr. Draxler (2004) SpeechRecorder – a Universal Platform Independent Multi-Channel Audio Recording Software, In. *Proc. of LREC*, pp. 559-562, Lisbon
- F. Schiel (2004) MAUS goes iterative, In. *Proc. of LREC*, pp. 1015-1018, Lisbon
- Chr. Draxler, K. Jänsch (2008) WikiSpeech – A Content Management System for Speech Databases, In. *Proc. of Interspeech*, pp. 1646-1649, Brisbane