

# Transkriptionsleitlinien

---

Eine Transkription ist eine orthographische Repräsentation einer Audiodatei. Für den VoiceAward-Korpus stellen wir die Transkription und darauf aufbauend eine Annotation von Dialogen mit automatischen Sprachdialogsystemen her. Jeder Dialog (= ein Anruf) liegt in einer eigenen Datei vor und wird einzeln transkribiert. Dialoge bestehen aus Turns. Jeder Sprecherwechsel (von Benutzer zu Dialogsystem oder umgekehrt) markiert den Beginn eines neuen Turns. Für die Transkription benutzen wir das frei erhältliche Programm *Transcriber*.

**Achtung:** Wenn etwas unklar oder besonders interessant ist, bitte immer die Stelle aufschreiben!

## Grundlagen der Benutzung von *Transcriber*

- Lade *Transcriber* herunter: <http://trans.sourceforge.net/> Installiere das Programm.
- Eine Bedienungsanleitung gibt es unter <http://trans.sourceforge.net/en/usermanUS.php>
- Um eine neue Sitzung zu beginnen, klicke auf das Menü „File/New trans“. Wähle die Audiodatei, die transkribiert werden soll.
- Speichere oft mit „File/Save“. Beim ersten Speichern einer neuen Transkriptionsdatei wird ein Dateiname vorgeschlagen. Dieser sollte nicht geändert werden!
- Wenn eine begonnene Transkription weiter bearbeitet werden soll, kann diese mit „File/Open trans“ geöffnet werden. Hierfür ist es notwendig, dass die Transkriptionsdatei den zur Audiodatei passenden Namen hat (siehe vorherigen Punkt).
- Um beim nächsten Starten von *Transcriber* dieselbe Transkription automatisch zu öffnen, speichere die Konfiguration vor dem Schließen mit „Options/Save configuration“.
- „Tab“ startet und pausiert das Abspielen des Audiosignals.
- Das Audiosignal kann mit der Maus, den Menüs oder mit Tastaturkürzeln manipuliert werden.

## Ablauf einer Transkription

Bei der Transkription soll Präzision und Schnelligkeit ausbalanciert werden. D.h., eine möglichst korrekte Transkription soll in möglichst kurzer Zeit hergestellt werden. Eine Transkriptionssitzung besteht aus drei Durchgängen, die im Folgenden erklärt sind.

### **Erster Durchgang: Grobe Segmentierung**

Zur Erleichterung der eigentlichen Transkription und zur Sprecherkennzeichnung muss das durchgängige Sprachsignal in mehrere Abschnitte segmentiert werden. Im ersten Durchgang geht es um eine grobe Einteilung, die genauen Abschnittsgrenzen können bei der Transkription im zweiten Durchgang korrigiert werden.

- Wähle eine Auflösung von maximal 30s für das Signal.
- Spiele das Signal ab.
- Während des Durchlaufs, setze eine Segmentierung (durch „Enter“):
  - bei einer größeren Sprechpause, die mit einem Satzende zusammenfällt
  - bei Sprecherwechsel
  - zu Beginn und Ende von nichtsprachlichen Signalen (z.B. Musik)

## Transkriptionsleitlinien

Eine neue Textzeile erscheint, und unter dem Audiosignal erscheint ein senkrechter Strich an der Position des Cursors (ein „Breakpoint“).

- Achte darauf, dass Segmentgrenzen innerhalb von Pausen liegen, und keine Sprache davon abgeschnitten wird.
- Speichere oft („File/Save“)

### Zweiter Durchgang: Transkription

- Wähle eine Auflösung von 10s für das Audiosignal.
- Gehe zum ersten untranskribierten Segment durch Klicken auf die erste Textzeile oder den Anfang des Signals.
- Höre den Abschnitt an (Menü „Signal/Replay segment“).
- Wenn nötig, passe die Abschnittsgrenzen durch Klicken mit der mittleren Maustaste und Ziehen an. Abschnittsgrenzen sollten immer im stummen Bereich liegen und nichts abschneiden!
- Wenn am Beginn dieses Abschnitts ein neuer Sprecher anfängt, markiere diesen Sprecherwechsel mit „Segmentation/Create turn“. Als Sprechernamen sollen „SDS“ für das Dialogsystem und „User“ für den menschlichen Anrufer benutzt werden. („Other“ = menschlicher Berater, anderer signifikanter Sprecher)
- Gib die Transkription über die Tastatur ein. Sie erscheint sowohl in der Textzeile als auch in der Segmentierung unter dem Audiosignal.
- Transkribiere rein nichtsprachliche Segmente (z.B. Musik) mit „#NS“. Benutze als Sprechernamen den Standardeintrag „(no speaker“.
- Wiederhole dies für alle verbleibenden Abschnitte.

### Dritter Durchgang: Verifikation

- Spiele die gesamte Datei erneut ab und lies die Transkription parallel. Das Abspielen kann mit „Tab“ sooft wie nötig gestoppt und wieder gestartet werden. Vor- und Rückspulen ist auch möglich.
- Transkribierter Text kann im Textfenster korrigiert werden.
- Sprechernamen werden mit dem Menü „Segmentation/Edit attributes“ geändert.
- Segmentgrenzen können durch Ziehen mit der mittleren Maustaste angepasst werden.

## Richtlinien

1. Transkribiere, was gehört wird, ohne grammatische Fehler zu korrigieren.
2. Benutze normale Groß- und Kleinschreibung.
3. Achte auf häufige Schreibfehler wie „dass“/„das“, usw.
4. Wenn ein Wort auch nach mehrmaligem Hören unverständlich ist, transkribiere, was Du hörst, und markiere das Wort mit Klammern, z.B.: „(fligop)“. Dies sollte jedoch so viel wie möglich vermieden werden.
5. Zahlen sollten ausgeschrieben werden, z.B.: „vierundvierzig“ statt „44“.
6. Wörter sollten ausgeschrieben werden (also „Straße“ statt „Str.“).
7. Einzelne Buchstaben werden mit einem nachfolgenden Unterstrich markiert: „A\_“.
8. Abkürzungen sollen so transkribiert werden, wie sie gesprochen werden. Also z.B. „U\_ S\_ A\_“ oder „Uno“.
9. Wenn ein Sprecher ein Wort nicht beendet, und es klar ist, um welches Wort es sich handelt, schreibe das Wort soweit aus, wie es gesprochen wurde, und füge einen Bindestrich an. Z.B.: „Also ich glau-“.
10. Verwende normale Interpunktion, aber spärlich. Nur Komma, Punkt und Fragezeichen sollen in Sätzen benutzt werden. Wenn Bindestriche innerhalb von Wörtern

vorkommen, darauf achten, dass kein Leerzeichen zwischen Bindestrich und Wort entsteht.

11. In Zahlenketten, setze keine Interpunktion zwischen die einzelnen Ziffern. Also z.B.: null drei null drei neun vier neun eins acht null fünf.
12. Überlappungen können entstehen, wenn der Benutzer das System unterbricht („Barge-In“), oder umgekehrt. Der Beginn der Überlappung wird als Beginn des neuen Segments (Turns) gekennzeichnet. In den Attributen des Turns kann als Merkmal „Überlappung“ („overlapping speech“) abgehakt werden, und dann sollte der zweite Sprecher ausgewählt werden (siehe Bedienungsanleitung von *Transcriber*).
13. Tastatureingaben des Benutzers werden durch „#KEY“ transkribiert (je eins pro Tastendruck).
14. Laute wie Räuspern, Husten, Lachen, etc., die vom System missverstanden werden könnten, werden durch „#“ repräsentiert.
15. Hintergrundgeräusche, z.B. das Hereinsprechen eines anderen Sprechers, werden nur dann transkribiert, wenn sie sehr laut sind und/oder so prominent, dass sie das System verwirren könnten. Dann werden sie als „#BG“ markiert.
16. Für Zögerlaute verwende „äh“ und „hm“.
17. Fremdsprachliche Wörter sollen mit „#FW“ markiert werden. Falls möglich, können die Wörter trotzdem transkribiert werden. Dann wird je ein „#FW“-Tag für den Anfang und das Ende des fremdsprachlichen Teils benutzt: „#FW For English, say English. #FW“ Ansonsten, also z.B. bei unbekanntenen Sprachen, kann einfach ein „#FW“-Tag zur Kennzeichnung des gesamten unverständlichen Abschnittes benutzt werden.

Bei Fragen wende dich bitte an Tatjana Scheffler (tatjana.scheffler@dfki.de oder 030-3949 1805).