

Jenny Winterscheid

## **Transkriptionsaufwand**

*Bearbeitungsstand: 29.05.2015*

„Da die Produktion gesprochener Sprache in mündlicher Interaktion ein schnell ablaufendes Geschehen“ darstellt (Selting 2001, 1059), müssen diese „flüchtige[n] Ereignisse“ für die Arbeit mit gesprochener Sprache zum einen „durch Aufzeichnung konserviert“ und anschließend transkribiert werden (Deppermann 2008, 21/39ff.).

Während die Erhebung der Daten zwar zeitaufwändig, aber doch noch relativ zügig zu bewältigen ist, handelt es sich bei der Transkription um eine langwierige Aufgabe, die oft unterschätzt wird. Damit Studierende, Promovierende sowie WissenschaftlerInnen, die Exposés, Arbeitspläne usw. aufstellen oder Zuschüsse für Projekte beantragen müssen, zukünftig über fundiertere Werte verfügen und damit realistischere Aussagen über den benötigten Aufwand (zeitliche wie personelle Ressourcen) treffen können, stelle ich hier die Ergebnisse einer Erhebung des Transkriptionsaufwandes im FOLK-Projekt vor.<sup>1</sup>

Da wir im FOLK-Projekt nur Minimaltranskripte anfertigen, können wir bisher auch nur eine Aussage über den Transkriptionsaufwand bei der Erstellung von Minimaltranskripten, ihre Korrektur und deren Normalisierung geben.

1. Transkription mit FOLKER .....	1
2. Ersteingabe .....	2
3. Korrektur.....	4
4. Normalisierung.....	4
5. Fazit.....	5

### **1. Transkription mit FOLKER**

Im Projekt Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK, <http://agd.ids-mannheim.de/folk.shtml>) erstellen wir Minimaltranskripte nach den GAT-2-Konventionen (Selting et al. 2009) und verwenden dabei den von Thomas Schmidt

---

<sup>1</sup> Diese Daten müssen natürlich an das jeweilige Projekt angepasst werden, bieten aber – allein schon wegen des unterschiedlichen Materials, das im FOLK-Projekt aufbereitet wird – bereits einige Anhaltspunkte, die als erste Orientierung für die Planung des Transkriptionsaufwandes von verschiedenem Datenmaterial herangezogen werden können.

entwickelten Transkriptionseditor FOLKER.<sup>2</sup> Zum Zwecke der computergestützten Transkription wurden diese Konventionen angepasst.<sup>3</sup> Dies ist in den GAT-2-Konventionen als mögliche notwendige Maßnahme mitberücksichtigt, wenn man zur Transkription einen Transkriptionseditor benutzt:

Zum Transkribieren mit Transkriptionsprogrammen (Editoren) können andere Eingabekonventionen notwendig sein, mit denen GAT 2 zwar kompatibel sein soll [...], die es aber nicht festlegen will und kann. (Selting et al. 2009, 354f.)

FOLKER bietet verschiedene Vorteile für die Aufbereitung von Gesprächen. Darunter fallen z.B. die Syntaxkontrolle für Minimal- und Basistranskripte nach GAT-2 sowie die Interoperabilität mit anderen Editoren. Die verschiedenen Ansichten (Segmente-, Partitur- und Beitrags-Ansicht), zwischen denen der Nutzer hin- und herschalten kann, unterstützen dabei eine präzise und effiziente Arbeit. Diese drei Ansichten kann man – wie dies auch im FOLK-Projekt praktiziert wird – für eine „sinnvoll[e] Abfolge von Arbeitsschritten beim Transkribieren“ (Schmidt/Schütte 2015, 3f.) einsetzen: Während eine Ersteingabe des Transkripts am besten in der Segmente-Ansicht stattfindet, sollten die Feinnotation der Simultanpassagen und eine erste Korrekturdurchsicht des Ersteingebenden danach in der Partitur-Ansicht vorgenommen werden. Da die zusammenhängenden Segmente eines Sprechers in der Beitrags-Ansicht zusammenhängend dargestellt werden, kann hier dann die Endkorrektur erfolgen. Diese Arbeitsschritte werden im Folgenden als Ersteingabe zusammengefasst behandelt.

## **2. Ersteingabe**

Um die Dauer des anfallenden Transkriptionsaufwandes abschätzen zu können, muss man verschiedene Faktoren berücksichtigen: So hängt die Zeit, die man generell für die Ersteingabe eines Transkripts veranschlagen muss, zum einen von der Länge der Aufnahme, deren Qualität sowie Komplexität und zum anderen von der Erfahrung der Transkribierenden mit dem Erstellen von Transkripten, dem Transkripteditor und den Aufnahmen bzw. den Sprechern ab.

---

<sup>2</sup> Der Transkriptionseditor FOLKER wird ausgehend von der AGD-Seite „<http://agd.ids-mannheim.de/folker.shtml>“ nach Registrierung zum Download angeboten. Das zugehörige Transkriptionshandbuch (Schmidt/Schütte 2015) ist auf dieser Seite verfügbar unter [http://agd.ids-mannheim.de/download/FOLKER-Transkriptionshandbuch\\_preview.pdf](http://agd.ids-mannheim.de/download/FOLKER-Transkriptionshandbuch_preview.pdf).

<sup>3</sup> Das cGAT-Handbuch, das sich vor allem an FOLKER-NutzerInnen oder an Transkribierende, die gerne die GAT-2-Konventionen an Transkript-/Höbeispielen nachvollziehen wollen, richtet, befindet sich in der Vorbereitung und wird in Kürze über die Seiten des AGD veröffentlicht werden.

Das ausschlaggebendste Kriterium ist dabei die Erfahrung der Transkribierenden. Wenn sie noch ungeübt in der Arbeit mit FOLKER und den GAT-2-Konventionen ist, benötigen sie für die Ersteingabe einer Aufnahme, die als durchschnittlich eingestuft werden kann, ca. 1 Stunde pro Minute. Nach einer gewissen Einarbeitungszeit, die etwa mit zwei Monaten anzusetzen ist, wird hingegen für eine Minute nur noch knapp eine halbe Stunde benötigt.

Die Erfahrung ist entscheidend:


Erfahrene Transkribierende sind ca. 2,5 mal schneller als unerfahrene.

Als Referenz wurde hier eine Arbeitsinteraktion mit vier SprecherInnen zugrunde gelegt, da die Transkription einer solchen Aufnahme als durchschnittliche Aufgabe gelten kann.

Sind an solch einer Interaktion nur 3 Sprecher beteiligt, benötigen erfahrene Transkribierende statt ca. 30 Stunden nur noch ca. 24 Stunden für eine einstündige Aufnahme. Sind hingegen mehr Sprecher beteiligt, benötigen Transkribierende mehr Zeit für die Transkription. Bei 5 Sprechern wären es bei einem einstündigen Meeting bereits ca. 48 Stunden.

Bei einem Alltagsgespräch mit 5 Sprechern benötigen erfahrene Transkribierende über 50 Stunden, während für die Transkription einer geregelten Kommunikationssituation, etwa einer Lehr-Lern-Interaktion, nur ca. 20 Stunden Arbeitszeit veranschlagt werden müssen. Das liegt vor allem daran, dass es z.B. in Alltagsgesprächen zumeist mehr Überlappungen und undeutliche Passagen gibt.

Beispielhaft wird dies an 5 verschiedenen Interaktionen dargelegt:

Institutionelle Interaktion    Alltagskommunikation	Lehr-Lern-Interaktion (Lehrer-Klasse)  Meeting (3 Sprecher) Meeting (4 Sprecher) Meeting (5 Sprecher)  Gespräch unter Freunden (5 Sprecher)	ca. 20 Stunden pro Stunde  ca. 24 Stunden pro Stunde ca. 30 Stunden pro Stunde ca. 48 Stunden pro Stunde  ca. 51 Stunden pro Stunde
--	---	---

Neben der Erfahrung der Transkribenten muss also auch das Datum selbst bei der Kalkulation berücksichtigt werden:

Je informeller die Aufnahmesituation ist und je mehr Sprecher beteiligt sind, desto länger dauert die Ersteingabe.

### **3. Korrektur**

Den Korrekturprozess kann man unterschiedlich aufwändig gestalten. So kann man bei einer Durchsicht Anmerkungen hinzufügen oder die Änderungen gleich vornehmen. Rückmeldungen kosten Zeit, können aber auch aus didaktischer Sicht sinnvoll sein und bei einem längerfristigen Projekt auch nach einer gewissen Zeit wiederum Zeit sparen! Deswegen sollte man sich bei der Planung bereits überlegen, ob Korrekturen vorgenommen werden sollen und wie elaboriert dies geschehen soll.

Im FOLK-Projekt werden Transkripte nach der Ersteingabe korrigiert, um möglichst gute Transkripte zu erhalten. Dabei geht es erst einmal darum, Fehler zu verbessern, aber auch um die Homogenität der Transkripte, die über die [Datenbank Gesprochenes Deutsch](#) in der Version 2.4 angeboten werden (sollen). Bei der Korrektur sichten andere Transkribierende als die Ersteingebenden das Transkript und überprüfen den transkribierten Wortlaut, die Segmentgrenzen, die Pausen etc. Bei diesem Schritt ist natürlich essentiell, wie exakt der/die Ersteingebende gearbeitet hat und wie ausführlich die Anmerkungen sein müssen, da diese ja später von dem jeweiligen Autor bzw. der Autorin eingearbeitet werden sollen. Bei Transkribierenden, die bereits mehrere Transkripte erstellt haben, reichen oft die im Projekt vereinbarten Kürzel aus,<sup>4</sup> während bei neuen Transkribierenden die Anmerkungen generell etwas detaillierter ausfallen, da noch Tipps für die Bearbeitung des Reparandums mitgeliefert werden. Das sieht man beispielweise an der Bearbeitungsdauer eines Transkripts des Interaktionstyps „Meeting“ sehr gut: Die Korrektur eines Transkripts einer einstündigen Arbeitsinteraktion mit sechs Gesprächsbeteiligten dauert ca. 10 Stunden. Der Korrekturdurchgang eines von einem/einer noch relativ unerfahrenen Transkribierenden erstellten Transkripts desselben Interaktionstyps, in welchem sogar nur vier Gesprächsbeteiligte miteinander interagieren, dauert hingegen ca. 12 Stunden.

---

<sup>4</sup> Die Korrekturkonventionen sowie eine Beschreibung des Verfahrens können bei Bedarf bei den Projektmitarbeitern erfragt werden.

Erneut gilt:

Der Korrekturprozess dauert länger, je informeller die Aufnahmesituation ist, je mehr Sprecher beteiligt sind und wenn der/die TranskribentIn noch relativ unerfahren ist.

#### **4. Normalisierung**

Neben der Transkription in literarischer Umschrift werden in FOLK auch normalisierte Transkripte angefertigt, um diese besser durchsuchbar zu machen und um eine optimale Basis für das [Part-of-Speech-Tagging](#) zu schaffen. Hierfür wird das mit FOLKER angefertigte Transkript mit dem ebenfalls von Thomas Schmidt entwickelten Normalisierungseditor OrthoNormal<sup>5</sup> automatisch normalisiert und dieser automatische Prozess dann manuell korrigiert. Die automatische Normalisierung basiert auf einem hinterlegten Lexikon und den bisherigen Normalisierungsvorgängen durch die Transkribierenden. Eine Durchsicht besteht darin, dass die automatische Normalisierung an bestimmten Stellen zurückgenommen oder abgeändert und bisher unterbliebene Normalisierungen vorgenommen werden müssen. Darüber hinaus können (und sollen) auch Transkriptionsfehler, die dem/der Transkribierenden bei dieser Durchsicht auffallen, im Transkript korrigiert werden. Somit sind sehr standardnahe Gespräche mit einem geringeren Zeitaufwand zu bearbeiten als dialektal gefärbte sowie gute Transkripte, bei denen keine Korrekturen vorgenommen werden müssen.

Durchschnittlich kann für diesen Arbeitsschritt etwa ein Arbeitsaufwand von 4,5 Stunden für ein einstündiges Transkript angesetzt werden.

Auch hier gilt wieder:

Unerfahrene Transkribierende brauchen deutlich mehr Zeit für die Normalisierung als erfahrene und auch hier ist die Beschaffenheit des Transkripts bzw. der Aufnahme entscheidend.

---

<sup>5</sup> Der Normalisierungseditor OrthoNormal wird bei dem Windows-Download von FOLKER automatisch mit heruntergeladen, für Mac kann er ausgehend von der AGD-Seite <http://agd.ids-mannheim.de/folker.shtml> nach Registrierung separat heruntergeladen werden. Die Normalisierungskonventionen werden in Kürze ebenfalls auf der Seite des AGD zum Download angeboten.

## **5. Fazit**

Die Transkriptionsarbeit wird häufig unterschätzt. Bei der Kalkulation des Transkriptionsaufwands muss deswegen eine Entscheidung getroffen werden, ob man wirklich die ganze Interaktion transkribieren muss und welchen Detaillierungsgrad und Qualität das Transkript aufweisen muss. Für viele Analysen reicht es völlig aus, ein Gesprächsinventar (vgl. Deppermann 2008, 32-35) zu erstellen, daraus Passagen für die Basistranskription auszuwählen oder das Minimaltranskript einer Interaktion anzufertigen und bestimmte Passagen anschließend nach den GAT-2-Basistranskriptkonventionen auszubauen.

Im Projekt FOLK erstellen wir ausschließlich Minimaltranskripte. Dafür können wir festhalten, dass erfahrene Transkribierende für die komplette Transkription (inklusive Korrektur) einer einstündigen Aufnahme bei einer als durchschnittlich eingestuften Aufnahme etwa 54 Stunden ansetzen müssen. Bei der Kalkulation des Transkriptionsaufwands ist die Beschaffenheit der Aufnahme – gemeint sind insbesondere der Interaktionstyp und die Sprecheranzahl, aber auch die technische Qualität der Aufnahme – zu berücksichtigen. Dass man mehr Zeit für die Transkription ansetzen muss, wenn die Anzahl der am Gespräch beteiligten Personen oder der Interaktionstyp variieren, kann man daran eindrucksvoll zeigen, dass erfahrene Transkribierende bei einer kompletten Transkription einer Spielinteraktion mit sieben Sprechern bereits ca. 76 Stunden aufwenden müssen.

Die Analyse des Arbeitsaufwands hat im Übrigen ergeben, dass bei allen Arbeitsschritten die Erfahrung der Transkribierenden mit der Erstellung von Transkripten und dem Transkripteditor eines der wichtigsten Kriterien für die Bearbeitungsdauer darstellt, da die Ersteingabe mitunter bis zu 2,5-mal länger dauern kann, aber auch die Korrektur der erstellten Transkripte aufwändiger ist.

Die vorgestellten Werte beziehen sich alle auf die Erstellung von Minimaltranskripten nach den GAT-2-Konventionen; wenn Basis- oder gar Feintranskripte angefertigt werden sollen, muss der Transkriptionsaufwand um ein Vielfaches höher angesetzt werden. Dagegen lassen sich einfache Abschriften mit deutlich geringerem Zeitaufwand erstellen.

## **Bibliographie**

- Deppermann, Arnulf (2008): Gespräche analysieren. Eine Einführung. Wiesbaden: Verlag für Sozialwissenschaften, 4. Auflage.
- Schmidt, Thomas / Schütte, Wilfried (2015): *FOLKER Transkriptionseditor für das „Forschungs- und Lehrkorpus gesprochenes Deutsch“ (FOLK). Transkriptionshandbuch. Version für das Preview der Version 1.2.* Bearbeitungsstand: 19.03.2015. Abrufbar unter [http://agd.ids-mannheim.de/download/FOLKER-Transkriptionshandbuch\\_preview.pdf](http://agd.ids-mannheim.de/download/FOLKER-Transkriptionshandbuch_preview.pdf) [Stand: 07.05.2015]
- Selting, Margret (2001): Probleme der Transkription verbalen und paraverbalen/prosodischen Verhaltens. In: HSK 16.2, 1059-1068.
- Selting, Margret et al. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion 10, S. 353-402 (im Internet unter <http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf>).