Projekt: Digital unterstützte Spontansprachanalyse (DigiSpon)

Kontakt: Pascale Schaller (pascale.schaller@unifr.ch), Ramona Rüegg (ramona.rueegg@unifr.ch)

Transkriptionsregeln (Stand März 2025)

1) Die Gespräche werden mit der Transkriptionssoftware f4[1] transkribiert, die automatisch Zeitmarken setzt (falls nicht voreingestellt, dann einstellen).

2) Die Gespräche werden vollständig transkribiert, und zwar sowohl die Äusserungen der Fachperson (FP) als auch die Äusserungen des Kindes (K).

3) Wurde im Gespräch ein Name anonymisiert (durch Stille ersetzt), wird diese Stelle mit NAME gekennzeichnet

4) Dialekt/Schweizerdeutsch: Es wird grundsätzlich Kleinschreibung verwendet. Majuskeln werden für die Markierung der Betonung verwendet:

   K: dä het de gsäit nÄi dä rägebogefisch

   Standarddeutsch/Hochdeutsch: Die Gross-Kleinschreibung orientiert sich an standardsprachlichen Konventionen:

   K: wir haben einen Weihnachtsbaum zuhause

5) Verständigungssignale, Wiederholungen, Wort- und Satzabbrüche sowie Überlappungen werden ebenfalls erfasst.

6) Interpunktion wird nicht verwendet, eine Ausnahme bildet die Verwendung des Fragezeichens zur Markierung einer Frage, die eindeutig als solche erkennbar ist:

   FP: hier wird ein Weihnachtsbaum geschmückt feiert ihr zuhause auch Weihnachten?

   K: ja

   FP: ja?

   K: hm {bejahend} mühsam

   FP: das findest du mühsam?

7) Die Orientierung am standardsprachlichen Schriftbild respektive an standardsprachlichen Konventionen gilt für die Transkription von bestimmten Wörtern mit hohem Wiedererkennungswert: *ok*

8) Pausen werden durch Punkte in Klammern markiert. Es werden drei Längen unterschieden:

   (.) = ca. 1 Sek. / (..) = ca. 2 Sek. / (...) = ca. 3 Sek.

9) Verständigungssignale des gerade nicht Sprechenden wie *mhm*, *aha*, *ja* werden nur dann nicht transkribiert, wenn sie während des Redebeitrags des Gegenübers geäussert werden.

[1] Zeitmarken werden von anderen Programmen (z.B. MaxQda) übernommen.

10) Verständigungssignale werden mit geschwungenen Klammern wie folgt transkribiert:

    `hm {bejahend}`

    `hm {verneinend}`

    `ähm`

    `ah, oh uh` mit nur einem Vokal und einem Dehnungs-h, unabhängig von der Länge des Lautes

11) beschreibende Ergänzungen werden ebenfalls in geschwungenen Klammern ergänzt:

    `FP: tuesch dù au gärn baschtle {flüstert}`

12) Unverständliche Wörter werden wie folgt mit eckigen Klammern markiert:

    `[xxx/yyy]` = Zwischen zwei Varianten kann nicht zweifelsfrei entschieden werden

    `[xxx]` = unsicher, ob es wirklich das Wort ist

    `[+++]` = ein Wort oder eine Äusserung ist überhaupt nicht verständlich (die drei Pluszeichen stehen dabei, anders als in GAT (Selting et al. o.J.), entweder für ein einzelnes vermutetes Wort oder für mehrere vermutete Wörter).

13) Überschneidungen werden mit zwei Querstrichen zu Beginn und am Schluss der entsprechenden Passagen markiert:

    `FP: der Weihnachtsbaum //ist grün//`

    `K: //ja manchmal ist bei uns// auch ein Weihnachtsbaum`

14) Mazes (gefüllte Pausen, falsche Starts, Wiederholungen und Neuformulierungen) werden mit runden Klammern markiert:

    `K: und dän (äh äh chötem äh) chÖmer ebe nöd jedes jaar für alli e gebùrtstagparty (öh öh öh) machä`

    Wenn die Wahl besteht, welche Wörter oder Phrasen markiert werden sollen, wird das früheste Vorkommen in Klammern gesetzt und das jeweils letzte Vorkommen des Wortes oder der Phrase als die erfolgreiche Produktion.

    Erfolgt eine Wiederholung absichtlich, um einen Sachverhalt besonders zu betonen, wird keine Maze markiert, denn Mazes sind Anzeichen von Formulierungsschwierigkeiten.

15) Einteilung in C-Units (Communication-Units): In Anlehnung an die Salt Transcription Conventions werden die Äusserungen in Communication Units (C-Units) segmentiert. Es wird empfohlen, vor der Transkription einer Sprachprobe die hier verlinkten Transkriptions-Konventionen von SALT genau zu lesen, da dort umfangreiche Hinweise zur korrekten Segmentierung gegeben werden, die an dieser Stelle nicht alle aufgenommen werden können.

    Salt Transcription Conventions:
    https://www.saltsoftware.com/pub/media/wysiwyg/tranaids/StandardConventions.pdf

Generell gilt: C-Units werden definiert als „ein unabhängiger Satz und seine Modifikatoren“ (Loban, 1976). Ein Satz, ob es sich um einen Hauptsatz oder einen Nebensatz handelt, ist eine Aussage, die sowohl ein Subjekt als auch ein Prädikat enthält. Hauptsätze können für sich allein stehen und in eine C-Einheit unterteilt werden. Nebensätze sind vom Hauptsatz abhängig, um einen Sinn zu ergeben. Sie können nicht für sich allein stehen oder vom Hauptsatz getrennt werden. Eine C-Einheit besteht also entweder aus einem Hauptsatz oder einem Hauptsatz mit einem oder mehreren Nebensätzen.

C-Units in Zusammenhang mit Ja-/Nein-Antworten:

a) Die Ja/Nein-Antwort wird von der nachfolgenden Äußerung getrennt, wenn darauf eine vollständige Äußerung/C-Unit folgt:

```
K: ja (..)

K: und oise wychtel hett de tAig mach für d roipli
#00:02:12-2#
```

b) Eine Ja/Nein-Antwort nicht als eigenständig segmentiert, wenn darauf eine unvollständige Äußerung/C-Unit folgt:

```
FP: Hast Du ein Haustier?

K: Ja einen Hund
```

16) Äusserungen auf Schweizerdeutsch werden wie folgt transkribiert: Die Dialekt-Schreibweise richtet sich nach dem Dieth’schen Prinzip: Sie ist möglichst nahe an der Lautung, orientiert sich aber gleichzeitig so weit wie möglich am gewohnten und dem Leser vertrauten (standardsprachlichen) Schriftbild:

a) Länge wird durch Doppelvokale (*Maa* ‚Mann’), Kürze durch einfache Vokale (*bache* ‚backen’) markiert.
b) Die Qualität der Vokale *i*, *u* und *ü* wird folgendermassen angegeben: offenes dumpfes *i* (*spine* ‚spinnen’, *hiiter* ‚heiter’) als *i*, geschlossenes helles *i* (*rytte* ‚reiten’, *fryy* ‚frei’) als *y*; offenes dumpfes *u* (*rùnd* ‚rund’, *tùùb* ‚wütend’) als *ù*, geschlossenes *u* (*brunn* ‚braun’, *stuune* ‚staunen’) als *u*; offenes *ü* (*nùmm* ‚nimm’, *Zùùg* ‚Züge’) als *ů*, geschlossenes *ü* (*Trüble* ‚Trauben’, *Züüg* ‚Zeug’) als *ü*.
c) Kontraktionen wie *han i ne* ‘hab ich ihn’, *isch ne* ‘ist ihnen’ werden an der Wortgrenze der standardsprachlichen Entsprechung segmentiert

17) St/sp werden orientiert am standardsprachlichen Schriftbild transkribiert: [ʃtɑːb̥] ‘Stab’ als `staab` (analog sp)

18) Die Transkripte werden im txt-Format gespeichert.