

## Mehrebenenannotation multimodaler Daten

Eva Reimer / Anna Valentine Ullrich  
Bianka Trevisan / Eva-Maria Jakobs

### 1. Einleitung

Die Untersuchung gesprochener Sprache (Spontansprache) war lange ein Forschungsbereich, der primär durch die Gesprächsforschung bedient wurde. Die Analysen erfolgten manuell. In den letzten Jahren wächst das Interesse daran, Methoden und Ansätze der Computerlinguistik und Texttechnologie zu nutzen, um Analysen von Spontansprache computerbasiert zu unterstützen. Die DFG hat dazu eine Handreichung mit Empfehlungen zu Datenstandards und Erhebungstools veröffentlicht (DFG 2015).

Der vorliegende Beitrag präsentiert einen komplexen theoretisch-methodischen Ansatz für die semi-automatische Annotation und Analyse von Spontansprache in Transkripten als Teil multimodaler Daten (Videos und Transkripte). Anhand von Videos und Transkripten professioneller Interaktionssituationen, in denen gesprochen, geschrieben und gezeichnet wird, steht die Beschreibung und Analyse von Modalitätsinterdependenzen (MID), dem Zusammenspiel der Modalitäten Sprechen, Schreiben und Visualisieren, im Fokus der Untersuchungen. Der Ausdruck *Annotation* bezieht sich im hier diskutierten Zusammenhang nicht auf die in der Gesprächsanalyse übliche Lesart (Transkriptionskonventionen zur Verdeutlichung gesprochener Sprache und ihre Handhabung, z.B. Annotation von Intonation oder Gleichzeitigkeit des Sprechens; Schneider 2002). Er bezeichnet vielmehr den Prozess (und das Ergebnis) der systematischen Anreicherung geschriebener oder gesprochener Sprache mit linguistischen Informationen, etwa für Zwecke der Mustererkennung (Ule/Hinrichs 2004:217). Die Informationen werden den Transkripten abstrahiert in Form von *Tags* (Klassen von Informationen) manuell oder automatisch hinzugefügt (Leech 1993:275).

Die Annotation erfolgt mit einer definierten Zielstellung, basierend auf einem Annotationsschema und zugehörigen Annotationsrichtlinien (vgl. Trevisan 2014: 70-71). Analyseleitend sind Fragen zu verschiedenen Typen von MID, MID-Akteuren und MID-Indikatoren sowie das Auftreten von MID in den untersuchten Interaktionstypen, z.B.

- In welchen Situationen treten MID-Typen auf?
- Welche MID-Typen treten auf – in welchem Interaktionstyp?
- Wer sind MID-Akteure?, Welche Indikatoren für MID treten auf?
- Was sind besonders starke oder nur vereinzelt auftretende Indikatoren?

Der vorliegende Untersuchungsgegenstand ist in besonderem Maße herausfordernd: Der multimodale Analysefokus (Modalitätsinterdependenzen, MID) erfordert einen breiten Annotationsansatz, um die Fülle relevanter Phänomene (z.B. Erfassung relevanter Objekte, Simultansprechen der Interaktionsbeteiligten) im Zusammenspiel von Sprechen, Schreiben und Visualisieren erfassen zu können. Im beschriebenen Mehrebenenannotationsansatz erfolgt die Annotation schichtweise, d.h. die Daten werden auf verschiedenen (Annotations-)Ebenen mit Infor-

mationen angereichert. Mit der Komplexität der Zielstellung wächst auch der Umfang des Annotationsschemas. Dies trifft in besonderem Maße auf den Gegenstand dieses Beitrages zu, die Annotation multimodaler Transkripte. Multimodale Transkripte enthalten neben den transkribierten Äußerungen Informationen zu praktischen und kommunikativen Handlungen der Interaktionsbeteiligten sowie Informationen zu in der Interaktion genutzten Objekten (Beispiel aus dem zugrundeliegenden Fallbeispiel: *PM1 hält B3 das Klemmbrett hin und zeigt darauf*). Sie erfordern die Definition und Einführung einer Vielzahl von Annotationsebenen, die neben sprachlichen auch kommunikative und praktische Handlungen im Sinne von Fiehler (1980) erfassen.

Nach einem Überblick über den Annotationsgegenstand und den Forschungszusammenhang (Kapitel 2) wird der Stand der Forschung zur Annotation von Spontansprache und Multimodalität dargestellt (Kapitel 3). Anschließend wird das Beschreibungsmodell für MID-Formen vorgestellt, für den der Ansatz entwickelt wurde (Kapitel 4). Darauf aufbauend wird die Datenbasis, auf die sich die Entwicklung des Annotationsansatzes stützt, beschrieben, gegenstandsspezifische Annotationsanforderungen aufgezeigt sowie das Annotationsschema erläutert und theoretisch verortet (Kapitel 5). Der Beitrag schließt mit einem Fazit und Ausblick (Kapitel 6). Das Annotationsschema wird vollständig im Anhang abgebildet.

## 2. Forschungskontext: Das Projekt ModiKo

Der hier präsentierte Annotationsansatz wurde in dem DFG-geförderten Projekt ModiKo "Sprechen – Schreiben – Visualisieren. Formen, Funktionen und Störungen von Modalitätsinterdependenzen (MID) und Konzeptgenesen in professionellen Interaktionssituationen"<sup>1</sup> für die Beschreibung (und spätere Analyse) von Modalitätsinterdependenzen (MID) entwickelt. *Modalitätsinterdependenz* wird definiert als das Zusammenspiel komplexer Ausdrucksressourcen in professionellen Interaktionssituationen (Ullrich et al. 2016). Die Interaktionsbeteiligten nutzen und kombinieren in der Interaktion verschiedene Ausdrucksressourcen wie z.B. Sprechen, Schreiben und Visualisieren. In der Kombination der Modalitäten entstehen Modalitäts-interdependenzen (MID): Die Interaktionsbeteiligten sprechen, machen sich parallel Notizen und zeichnen Sachverhalte, etwa zur Verständnissicherung, auf. Ziel des Projektes ist die Entwicklung von Modellen, Methoden und Tools, die es erlauben, MID systematisch zu beschreiben und zu analysieren.

Das Forschungsprogramm basiert auf gesprächsanalytischen Ansätzen, die gegenstandsbezogen erweitert werden durch korpus- und texttechnologische sowie computerlinguistische Ansätze. Teil des Projektes ist die Entwicklung eines Annotationstools für heterogene Datenbestände, mit dem Datenformate über mehrere Ebenen annotiert sowie Annotationen datenformatübergreifend in Bezug gesetzt und in ihrem Bezug dargestellt werden können.

Die Arbeiten in ModiKo stützen sich auf ein Fallbeispiel eines bestimmten Situationstyps: die professionelle Erhebung und -modellierung des Ist-Zustandes eines industriellen Fertigungsprozesses in einem Unternehmen. Das Fallbeispiel

<sup>1</sup> Das DFG-Projekt (GZ: JA 1172/3-1) findet in Kooperation mit Reinhard Fiehler und Thomas Schmidt vom IDS Mannheim statt.

umfasst drei zeitversetzt realisierte Interaktionstypen und ihnen zuzuordnende Handlungsschemata mit Rollen und kommunikativen und interaktionstyp-spezifischen Aufgaben der Interaktionsbeteiligten (Jakobs et al. 2011). Alle drei Interaktionstypen enthalten Modalitätsinterdependenzen. In Interaktionstyp 1 (Handlungsschema: *Interview als Mittel industrieller Prozesserhebung*) interviewt ein externer Experte (Prozessmodellierer PM1) Unternehmensmitarbeiter zu ihrer Arbeit im Fertigungsprozess.<sup>2</sup> Interaktionsbegleitend macht er sich Notizen und/oder fertigt Skizzen an. Die Notizen und Skizzen werden situationsversetzt (Interaktionstyp 2) von den Modellierern genutzt, um gemeinsam die in Interaktionstyp 1 mündlich erhobenen Informationen zu rekonstruieren, auf wesentliche Inhalte zu reduzieren und diese grafisch-symbolisch zu verdauern (Handlungsschema: *Vorläufige Prozessmodellierung durch externe Experten*). In Interaktionstyp 3 überführen die Modellierer interaktiv die Teilmodellierungen in eine finale Modellierung des Fertigungsprozesses, d.h. eine grafisch-symbolische Darstellung des Gesamtprozesses (Handlungsschema: *Übergabe an Dritte und finale Prozessmodellierung*) (vgl. Abb. 1).

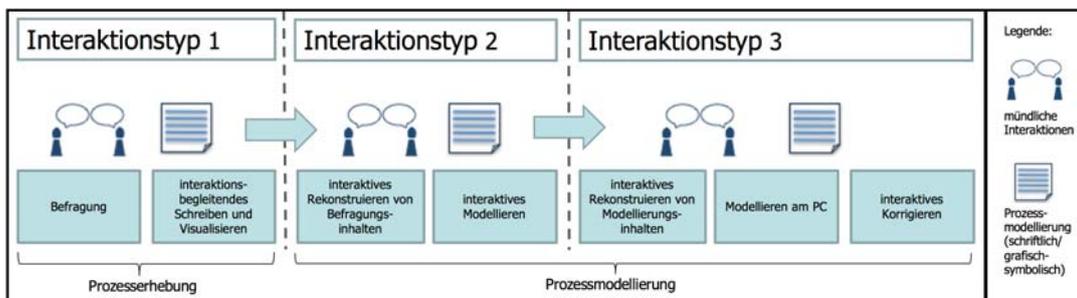


Abb. 1: Interaktionstypen des Fallbeispiels (in Anlehnung an Jakobs et al. 2011: 235)

In der ersten Projektphase konzentriert sich das Interesse auf die Erfassung, Beschreibung und Analyse von MID-Formen. In einem zweiten Schritt wird der Blick erweitert auf die Analyse von MID-Funktionen – bezogen auf einzelne Interaktionstypen wie auch interaktionstypübergreifend. Der Annotationsansatz unterstützt dies, indem er ermöglicht, Muster über verschiedene Datensätze hinweg zu betrachten und zu identifizieren.

### 3. Ansätze zur Annotation von Spontansprache und Multimodalität

Für die Annotation geschriebener wie gesprochener Sprache gibt es eine Reihe von Methoden, die für die Zwecke des Projektes zumindest in Ansätzen übernommen und/oder adaptiert werden. Ähnliches gilt für die Annotation multimodaler Daten, insbesondere die Annotation von Video- und Audiodaten. Die Ansätze werden im Folgenden kurz beschrieben.

Für die (automatische) Annotation geschriebener Sprache (Textkorpora) existieren verschiedene Ansätze, Korpora, Tools und Verfahren. Ein Standardtool für das Deutsche ist der TreeTagger (Schmid 1995), der die automatische Auszeich-

<sup>2</sup> Alle im Untersuchungskorpus betrachteten Akteure (Prozessmodellierer, interviewte Mitarbeiter) sind männlichen Geschlechts. Aus diesem Grunde wird durchgängig die maskuline Form verwendet.

nung von Texten mit morphosyntaktischen Kategorien (*Part of Speech*-Tags oder *PoS*-Tags, z.B. VVFIN = finites Verb) nach dem STTS-Tagset (Schiller et al. 1999) erlaubt. Die derart vergebenen PoS-Tags bilden die Grundlage und Voraussetzung für nachfolgende Annotationen. Untersuchungen zeigen jedoch, dass der TreeTagger bei der Verarbeitung nicht-standardisierter Texte wie Webkommentar (ca. 91% Genauigkeit, Neunerdt et al. 2013a:146) und Spontansprache (gesprochene Sprache) (ca. 81% Genauigkeit, Westphal/Schmidt 2013:149) eine geringe Performanz aufweist. Speziell für das PoS-Tagging von nicht-standardisierten Texten aus dem Internet ist der *WebTagger* entwickelt worden (Neunerdt et al. 2013b). Der WebTagger ist eine modifizierte Version des Standard TreeTaggers und erreicht für Webkommentare eine Genauigkeit von 94% (Neunerdt et al. 2013b:68). Trevisan (2014) entwickelt einen Mehrebenenannotationsansatz für die Annotation sprachlichen Bewertens in Blogkommentaren. Der Ansatz baut auf den PoS-Tags des WebTaggers auf und erlaubt die Auszeichnung von Äußerungsanteilen auf verschiedenen linguistischen Ebenen (z.B. auf graphematischer, lexikalisch-semantischer oder pragmatischer Ebene). Die linguistische Mehrebenenannotation wird mit dem Tool EXMARaLDA umgesetzt (Trevisan 2014:126; Schmidt/Wörner 2005). Das Tool *AutoAnnotator* erlaubt die semi-automatische Auszeichnung von Argumenten (Argumentation Mining) und baut auf dem Mehrebenenannotationsansatz von Trevisan (2014) auf. Ziel der Annotation ist die Identifikation von Argument-anzeigenden sprachlichen Ausdrücken und Mustern (Trevisan et al. 2014). Für das Deutsche gibt es eine Reihe weiterer Ansätze zur Untersuchung und Annotation von Nichtstandardsprache, die auf Phänomene wie Ad-hoc-Schreiben oder die Verwendung von Akronymen und Emoticons abzielen (vgl. Eller/Hirschmann 2014). Das Dortmunder Chat-Korpus ist ein Korpus zur Untersuchung internetbasierter Kommunikation und fokussiert die Sprachverwendung und sprachliche Variationen in der deutschsprachigen Chat-Kommunikation (vgl. Beißwenger 2013). Annotiert werden z.B. Nickname, Emoticons, Asteriken und Adressierungen in den Chat-Nachrichten (Beißwenger 2013:9).

Die Annotation gesprochener Sprache ist besonders herausfordernd: Ungenauigkeiten der morphosyntaktischen Annotation betreffen insbesondere die korrekte Auszeichnung von Partikeln und Interjektionen, Pronomen, Verben und "Nicht-Wörtern" (Genauigkeit: 81,16%, Westpfahl/Schmidt 2013). Am IDS-Mannheim wurde deshalb – basierend auf dem STTS-Tagset für geschriebene Sprache – ein erweitertes kohärentes Tagset für die morphosyntaktische Annotation gesprochener Sprache entwickelt, das die Besonderheiten von Spontansprache berücksichtigt (Westpfahl 2014). Es erweitert das STTS-Tagset durch neun Subkategorien, z.B. Subkategorien zur Auszeichnung von Onomatopoetika (NGONO) oder Intensivierungspartikeln (PTKIFG). Der TreeTagger wurde mit dem erweiterten Tagset auf dem Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK) trainiert und ein Goldstandard (FOLK-Gold) erstellt (Westpfahl/Schmidt 2016).

Im deutschsprachigen Raum gibt es neben dem FOLK-Korpus weitere annotierte Korpora gesprochener Sprache, diese wurden jedoch nicht mit spezifischen Tagsets bezogen auf die Untersuchung von Modalitätsinterdependenzen annotiert: Die Tübinger Baubank des Deutschen/Spontansprache (TüBa-D/S) ist ein syntaktisch annotiertes Korpus von ca. 360.000 Wörtern. Für die Annotation wurde

das STTS-Tagset in seiner ursprünglichen Form genutzt (Hinrichs et al. 2010; Telljohann et al. 2013).

Das KiezDeutsch-Korpus (KiDKo) ist ein transkribiertes, orthografisch normalisiertes Korpus spontaner Unterhaltungen von Jugendlichen im multi- und monoethnischen urbanen Raum (Wiese et al. 2012; Rehbein et al. 2014). Die Transkription erfolgte nach GAT 2 (Selting et al. 2009) in EXMARaLDA (Schmidt/Wörner 2005). Die Transkripte sind mit den Audiodateien in EXMARaLDA verknüpft (Rehbein/Schalowski 2013).

Neben Tools für die Annotation geschriebener und gesprochener Sprache gibt es Tools für die Annotation multimodaler Phänomene,<sup>3</sup> die die Annotation von Audio- und Videodaten ermöglichen: z.B. ANVIL (Kipp 2014), ELAN (Sloetjes et al. 2011) und Transana (Halverson 2012; Schwab 2006). ANVIL ist ein Annotationstool für Videodaten, das vor allem auf die Annotation von Gesten oder Kopfbewegungen abzielt. ELAN unterstützt die Verknüpfung und Annotation von Videos und Audiodateien in einem Screen. Beide Tools ermöglichen Annotationen auf mehrere Ebenen (z.B. für verschiedene Gesten). Eine Verknüpfung von Videodaten mit Text (z.B. Gesprächstranskripten) und/oder mit Scans (z.B. der in einer Interaktionssituation erstellten Notizen) ist nicht vorgesehen. Transana dient vorrangig der Transkription multimodaler Daten und erlaubt nur in geringem Umfang Analysen von Video- und Audiodaten. Eine Suchfunktion fehlt; umfangreiche Annotationen, beispielsweise auf morphosyntaktischer Ebene, sind nicht möglich.

Die genannten Ansätze erfassen Teilaspekte des in diesem Beitrag fokussierten Annotationsgegenstandes (Modalitätsinterdependenzen). Sie erlauben jedoch nicht, Modalitätsinterdependenzen in der Vielfalt der damit verbundenen Phänomene abzubilden. Insbesondere gibt es kein Tool, das die Annotation multimodaler Phänomene auf Grundlage eines multimodalen Transkripts ermöglicht.

#### 4. Beschreibungsmodell für MID-Formen

Die Entwicklung des Annotationschemas stützt sich auf einen Modellierungsansatz für MID-Formen, der datenbasiert in ModiKo entwickelt wurde. Da die Kenntnis des Modells grundlegend ist für den Nachvollzug des Annotationsansatzes bzw. -schemas, wird es im Folgenden kurz vorgestellt (ausführlich dazu Ullrich et al. im Druck).

Das pragmatisch ausgerichtete Modell betrachtet Modalitätsinterdependenzen aus zwei Perspektiven: makro- und mikroperspektivisch. Makroperspektivisch ist die mehrfache "Rahmung" des Auftretens von MID in Interaktionen relevant (vgl. Abb. 2). Sie ergibt sich durch den allgemeinen Situationstyp (im Fallbeispiel die professionelle Situation der industriellen Prozesserhebung und -modellierung in einem Unternehmen), den Interaktionstyp (hier: Interaktionstyp 1, 2 und 3) und die dazugehörigen Handlungsschemata mit Rollen und Aufgaben der Interaktionsbeteiligten.

Mikroperspektivisch sind MID-Formen diagrammatisch erfass- und beschreibbar: horizontal in der zeitlichen Abfolge von MID und speziellen, optional vor

<sup>3</sup> Einen Überblick über Tools zur Annotation multimodaler Daten bieten Duncan et al. (2013) und Wittenburg (2008).

und nach MID auftretenden Aktivitäten, vertikal über vier Betrachtungsdimensionen. Die Analyse der Falldaten zeigt, dass MID durch projektionsindizierende Aktivitäten und/oder retrospektive Verdeutlichungen der Interaktionsbeteiligten gerahmt auftreten können. Dies impliziert eine sequenzorientierte Darstellung und Betrachtung von MID (Achse *zeitliche Abfolge*).

Die Betrachtungsdimensionen Modality-taking, Manifestation, Akteur und interaktionaler Effekt erfassen dagegen Parameter der Interaktionssituation, die sich in der Datenanalyse als relevant für die Beschreibung von MID-Formen erwiesen haben (Achse *interaktionale Betrachtungsdimension*). Sie werden im Folgenden kurz beschrieben.

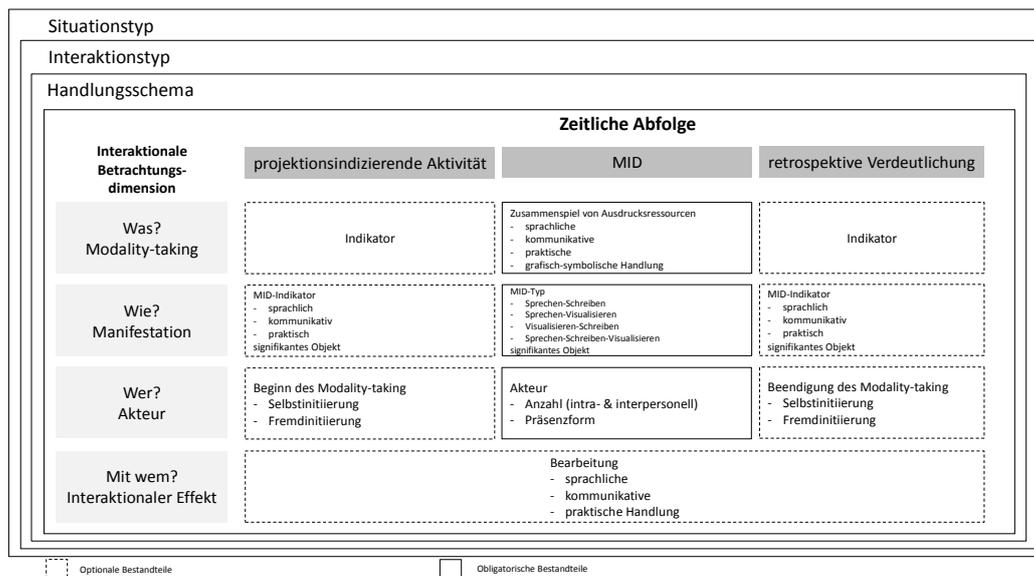


Abb. 2: Beschreibungsmodell für MID-Formen (Ullrich et al. im Druck)

Die Dimension *Modality-taking* erfasst, was bei einer MID geschieht, d.h. wie die Interaktionsbeteiligten das von ihnen in der Interaktionssituation genutzte Modalitätsspektrum erweitern (z.B. Sprechen durch Schreiben). Die situationspezifische Änderung des Modalitätsspektrums wird als *Modality-taking* bezeichnet (Ullrich et al. 2016:76). MID entstehen im Zusammenspiel sprachlicher, kommunikativer, praktischer und/oder grafisch-symbolischer Handlungen in Folge von Modality-taking. Indikatoren können prospektiv auf das Auftreten einer MID verweisen (Beispiel: mündliches Hinweisen auf eine Schreibhandlung: *das schreib ich mir als beispiel schon mal hier hin*). Nach dem Auftreten einer MID finden sich teilweise Hinweise auf die Beendigung einer MID (Beispiel: mündliches Hinweisen auf die Beendigung einer Schreibhandlung: PM1 notiert "Labels kommen von Gestell". PM1: *labels (-) kommen (-) von (-) gestell (-) muss ich sonst noch was zu dem prozess wissen (-)*).

Die Dimension *Manifestation* erfasst die Realisierung einer MID, die sich einem MID-Typ zuordnen lässt (spezifische Konstellationen von MID). Fallbezogen werden vier Typen von MID unterschieden: (1) *Sprechen-Schreiben*, (2) *Sprechen-Visualisieren*, (3) *Visualisieren-Schreiben* und (4) *Sprechen-Schrei-*

*ben-Visualisieren*.<sup>4</sup> Das Auftreten von MID kann gebunden sein an Objekte (z.B. Klemmbrett und Stift), die aufgabenbezogen signifikant werden.

Die Dimension *Akteur* erfasst die Akteure einer MID – ihre Anzahl (intra- und interpersonelle MID) sowie ihre Präsenzform (z.B. Sitzen am Tisch im Büro versus Stehen in der Produktionshalle). Akteurszentriert wird unterschieden, ob das Modality-taking (Änderung des genutzten Modalitätsspektrums) selbstinitiiert durch einen Akteur erfolgt und von ihm als solches thematisiert wird (Beispiel: *ich muss jetzt doch noch maln*) oder fremdinitiiert – ein Akteur fordert andere Interaktionsbeteiligte dazu auf (Beispiel: *ich würd sie einfach bitten*). In beiden Fällen handelt es sich um projektionsindizierende Aktivitäten. Analog finden sich im Korpus selbst- und fremdinitiierte Beendigungshandlungen als retrospektive Verdeutlichung einer MID (Beispiel: selbstinitiierte Beendigung in Interaktionstyp 1: Sprechpause des schreibenden Interviewers und anschließender Themenwechsel *mhm okay das muss ich mir mal grad aufschreiben (-)* ((PM1 notiert "10 Tressen!")) *ist bei der xkugel was besonderes zu beachten (-)*).

Die Dimension *Interaktionaler Effekt* im Kontext von MID erfasst, wie die Interaktionsbeteiligten mit Modalitätsänderungen ihrer Gesprächspartner (z.B. Wechsel von Schreiben zu Sprechen) umgehen. Die Aktivitäten im Interaktionsgeschehen umfassen sprachliche, kommunikative und praktische Handlungen (Beispiele: Sprechpause des Befragten während der Schreibphase des Interviewers oder Kooperativität anzeigende Side-by-Side-Stellung des Befragten neben dem sich Notizen machenden Interviewer und Blickwendung auf dessen Schreibblock).

## 5. Mehrebenenannotation

Der Mehrebenenannotationsansatz dient der systematischen Beschreibung von MID-Formen für folgende Analyseschritte, z.B. die Analyse der Musterhaftigkeit von MID und die Betrachtung von MID-Funktionen.

Die Entwicklung erfolgte daten- wie modellgeleitet. Im Folgenden wird die Datenbasis vorgestellt, anhand derer das Schema iterativ erprobt und weiterentwickelt wurde (Kapitel 5.1) sowie die gegenstandsspezifisch notwendigen Schritte der Datenaufbereitung und der automatischen Annotation beschrieben (Kapitel 5.2). In Kapitel 5.3 wird das Mehrebenenannotationsschema vorgestellt und die Annotationsebenen im Beschreibungsmodell für MID-Formen verortet.

### 5.1. Datenbasis

Für die Entwicklung des Annotationsansatzes wurde aus dem Gesamtdatenbestand des Projektes ein Pilotkorpus sowie – für die weitere Erprobung – ein Goldstandard gebildet. Die Ausgangsdatenbasis des Projektes bilden Primär- und Sekundärdokumente.<sup>5</sup> *Primärdokumente* umfassen die Videodateien der im Fallbeispiel erhobenen Interaktionen (548 Minuten, Format .mpeg) sowie die Scans der

<sup>4</sup> Die Bezeichnungen sind Kopulakomposita, sie geben keine Hinweise auf die Art und Weise des Auftretens der beteiligten Modalitäten (simultan oder sequenziell; ihre Reihenfolge).

<sup>5</sup> Wir danken in diesem Zusammenhang Reinhold Schmitt und Reinhard Fiehler für produktive, inspirierende Diskussionen.

interaktionsbegleitend erstellten Notizen und Skizzen (89 Seiten, Format .jpeg). Diese Daten wurden bereits in einem Vorgängerprojekt erhoben (IMIP: Interdisziplinäre Methoden industrieller Prozessmodellierung, BMBF, 2008-2011; Jakobs et. al 2011) und nach GAT 2 transkribiert. Diese Transkripte (266 Seiten, Format .doc) sind Grundlage der Sekundärdokumente. *Sekundärdokumente* sind multimodale Transkripte der Videodaten – *multimodal* meint hier die Anreicherung mit Informationen zu praktischen und kommunikativen Handlungen der Interaktionsbeteiligten sowie zu den in der Interaktion genutzten Objekten (z.B. *PM1 zeigt mit dem Stift auf das bereits Notierte; PM1 nimmt ein Klemmbrett vom Boden auf*). Durch die Mehrebenenannotation der Transkripte entstehen *Tertiärdokumente* (Reimer et al. 2015; Trevisan et al. 2016).

Die Auswahl der Teilkorpora (Pilotkorpus und Goldstandard) erfolgte in zwei Schritten:

- (1) Die Auswahl der Datensätze orientierte sich an "Interaktionsketten". Ausgewählt wurden Datensätze zu Interaktionen, die sich inhaltlich aufeinander beziehen und in der Zeit nacheinander in Interaktionstyp 1, 2 und 3 auftreten (vgl. Abb. 1). Beispiel: Interview des Prozessmodellierers PM1 mit einem Lageristen (Interaktionstyp 1), zeitversetztes Bearbeiten der Interviewinhalte durch PM1 und PM2 (Interaktionstyp 2), zeitversetztes kooperatives Integrieren der Ergebnisse in die Gesamtmodellierung (Interaktionstyp 3). Wenn ein und derselbe Gegenstand (z.B. LKW entladen) mehrfach in Interaktionen des gleichen Typs (z.B. Interaktionstyp 3) thematisiert wird, nimmt die Anzahl von Datensätzen pro Interaktionskette zu.
- (2) Die im ersten Schritt gesichtete Datenbasis wurde reduziert auf Interaktionsketten, die besonders reich sind bezogen auf das Auftreten von MID-Formen.

Für Zwecke der methodischen Entwicklung wurde das Pilotkorpus auf eine Interaktionskette reduziert (9.075 Token); es umfasst vier Datensätze (je einen aus Interaktionstyp 1 und 2 sowie zwei zu Interaktionstyp 3). Für den Goldstandard wurde das Pilotkorpus um vier Interaktionsketten erweitert. Der Goldstandard (33.185 Token) umfasst fünf Interaktionsketten mit 19 Datensätzen. Tabelle 1 bietet einen Überblick über die Datenbasis:

	Anzahl Token	Anzahl Interaktionsketten	Datensätze je Interaktionstyp			
			1	2	3	$\Sigma$
Pilotkorpus	9.075	1	1	1	2	4
Goldstandard	33.185	5	5	5	9	19
Gesamtkorpus	70.950	10	11	11	33	55

Tab. 1: ModiKo-Datenbasis

## 5.2. Datenaufbereitung und automatische Annotation

Die Umsetzung des Mehrebenenannotationsansatzes erfolgte in EXMARaLDA. Die transkribierten Äußerungen werden fortlaufend in einer Zeile dargestellt und durch Annotationsebenen (Zeilen) ergänzt. Anders als in Ansätzen zur Annotation

gesprochener Sprache können so neben morphosyntaktischen Kategorien Phänomene wie genutzte Ausdrucksressourcen (z.B. Sprechen oder Schreiben) oder Objekte (z.B. Stift) in separaten Ebenen annotiert werden.

Die Datenaufbereitung umfasst zwei Schritte: (1) Tokenisierung und Tagging der Daten und die Überführung der GAT 2-Transkripte in EXMARaLDA, (2) die automatische Annotation von Pausen, Wortabbrüchen und Simultansprechen. In beiden Schritten ergeben sich gegenstandsspezifische Herausforderungen.

Bei der *Tokenisierung* wird der Text in Segmente (Token) zerlegt; in der Regel umfasst ein Token ein Wort. Beim *Tagging* werden die Token morphosyntaktisch annotiert, d.h. mit PoS-Tags ausgezeichnet. Das Tagging erfolgt mit dem Tool TreeTagger. Wer sich – wie im Falle von ModiKo – dem Analysebereich Spontansprache aus der Perspektive der automatischen Analyse schriftsprachlich konzipierter und realisierter Texte (standardisierte Äußerungen, z.B. Zeitungstexte) nähert, muss bei der Tokenisierung und dem Tagging "umdenken", d.h. bei der Annotation und Toolgestaltung Anforderungen umsetzen, die sich aus Merkmalen des Gesprächstranskripts und der Spontansprache ergeben. Diese werden im Folgenden erläutert.

Zu den Merkmalen von Gesprächstranskripten gehört, dass sprachliche Äußerungen in Kleinschreibung sowie unter Verzicht auf Interpunktion erfasst werden. Interpunktionszeichen werden abweichend von der Schriftsprache für die Notation von Phänomenen wie Pausen und Verschleifungen genutzt. Die Vorgehensweise ist dem Gesprächsanalytiker vertraut – für Annotations- und Analyse-Tools, die für schriftsprachliche Texte entwickelt wurden, ist sie ungewöhnlich und erfordert erhebliche Anpassungen. Tools wie die Standardversion des TreeTaggers würden z.B. das Pausenzeichen (.) in drei Einheiten tokenisieren, d.h. es nicht als *eine* Einheit erkennen. Die Tools würden das Pausenzeichen auch nicht als Pausenzeichen interpretieren, da sie für standardisierte Schriftsprache trainiert sind (vgl. Neunerdt et al. 2013a; 2013b). Die adäquate Erfassung von Pausen ist jedoch für ModiKo gegenstandsbedingt relevant: Eine Annahme des Projekts ist, dass Pausen auf Modalitätsinterdependenzen hinweisen können und deshalb analyserelevant sind (Ullrich et al. im Druck). In ModiKo wurden deshalb Regeln entwickelt, die dem Annotationstool ermöglichen, im Transkript Interpunktionszeichen funktional eindeutig zu erkennen und auszuzeichnen. Sie beziehen sich auf folgende Transkriptionskonventionen (in Anlehnung an Selting et al. 2009; Jakobs et al. 2011):

ich/	Äußerungsabbruch
(.)	Mikropause
(-) (--) (---)	Pausen 1 bis 3 Sekunden
(8)	Pausen in Sekunden (Sekunden)
[ ]	Überlappungen und Simultansprechen.

Weitere Herausforderungen für Tools, die für Schriftsprache trainiert wurden, betreffen die Erfassung von Charakteristika der *Spontansprache* wie das gehäufte Auftreten von onomatopoetischen Ausdrücken (z.B. *au*), Inflektiven (z.B. *seufz*), Unterbrechungen, Wiederholungen und Reparaturen, Simultansprechen zwischen zwei oder mehr Sprechern, gehäuftes Auftreten von Partikeln (z.B. *sehr*), Interjektionen (z.B. *aha*, *naja*) und Responsiven (z.B. *ja*) sowie von verkürztem Spre-

chen, Verschleifungen und Tilgungen (z.B. *ich schreib das mal hier rein ähm*). Die genannten Phänomene bedingen neue Regeln, z.B. der Lemmatisierung (Rückführung von Wörtern bzw. Token auf ihre Grundform). Probleme bei der Lemmatisierung ergeben sich durch Kleinschreibung, starke Verkürzung von Wörtern oder wortverändernde Tilgungen (etwa bei Äußerungsabbrüchen, die bis zur Unkenntlichkeit von Wörtern führen). Tools wie die Standardversion des TreeTaggers zeigen eine geringe Performanz bei der Aufbereitung nicht-standardisierter Sprache; sie müssen speziell für die Anforderungen transkribierter Spontansprache trainiert werden. Die beschriebenen Herausforderungen werden in dem Annotationsansatz von Westpfahl (2014) berücksichtigt – allerdings begrenzt auf die morphosyntaktische Ebene. Die Komplexität der in ModiKo betrachteten multimodalen Transkripte erfordert einen deutlich breiteren Ansatz für die Erfassung von Phänomenen wie die Nutzung von Objekten in der Interaktion (z.B. signifikante Objekte wie Schreibblock, Stift) oder MID-Indikatoren. Dies ist möglich durch die Mehrebenenannotation. Für die Umsetzung wurde die Standardversion des TreeTaggers entsprechend modifiziert (auf Spontansprache trainiert; vgl. hierzu auch Westpfahl/Schmidt 2013, 2016; Westpfahl 2014).<sup>6</sup>

In EXMARaLDA werden die Sprechanteile der verschiedenen Sprecher parallel in Zeilen untereinander dargestellt und die Ebenen gegenstandsspezifisch belegt (vgl. Abb. 3 unten). Bei der Überführung der im Vorprojekt entstandenen Transkripte in EXMARaLDA (Sekundärdokumente) wurden die Zeitmarken anhand der Videos präzisiert und mit den Beschreibungen kommunikativer und praktischer Handlungen aligniert.

	838 [04:16.8]	839	840 [04:17.4]	841
PM1 [v]			telefonisch dann	
PM1 [tok]			telefonisch	dann
PM1 [pos]			ADJD	ADV
PM1 [lemma]			telefonisch	dann
PM1 [nv]			PM1 notiert „QM informiert“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.	
B3 [v]	(-) genau			
B3 [tok]	(-) genau			
B3 [pos]	XY	NGIRR		
B3 [lemma]	(-) genau			
B3 [nv]				

Abb. 3: Sekundärdokumente in EXMARaLDA  
(1a\_05\_prozesserhebung\_lagerist, Token 838-841)

In der obersten Zeile werden die Token nummeriert (in Abb. 3 Token 838-841) sowie Zeitmarken zugeordnet (z.B. Token 840 der Zeitmarke [04:17.4]). Die Benennung der Ebenen ergibt sich aus dem Sprecherkürzel und dem Kürzel für den in der Ebene notierten Inhalt. Jedem Sprecher (hier: PM1 und B3) werden fünf Zeilen zugeordnet: Die erste Zeile [v] erfasst die transkribierte Äußerung, die zweite Zeile [tok] die Token eines Äußerungsteils. Jedes Token wird in der dritten Zeile [pos] morphosyntaktisch beschrieben (PoS-Tag); in der vierten Zeile [lemma] erscheint das jeweilige Lemma. Die fünfte Zeile [nv] erfasst zeitmarkenkonform Beschreibungen kommunikativer und praktischer Handlungen des Sprechers (z.B. Gestik oder Interaktion mit Objekten).

<sup>6</sup> Für die Umsetzung, Änderungen und Anpassungen von Tokenisierung und Tagging danken wir Thomas Schmidt.

Diese Sekundärdokumente in EXMARaLDA bilden die Basis für die händische und automatische Mehrebenenannotation (Tertiärdokumente) zur Erfassung und Beschreibung von Modalitätsinterdependenzen für die spätere Analyse. Die analyserelevanten Phänomene *Wortabbrüche*, *Pausen* und *Simultansprechen* können automatisch ausgezeichnet werden, da die Transkription dieser Phänomene Regeln (Transkriptionskonventionen in Anlehnung an Selting et al. 2009; Jakobs et al. 2011) folgt.<sup>7</sup> Die automatische Annotation erfolgt mit dem Tool *AutoAnnotator* (Trevisan et al. 2014). Das Tool erlaubt die automatische Annotation regelbasiert beschreibbarer und gering interpretativer Annotationsebenen. Die Transkriptionsregeln für die genannten Phänomene (z.B. ein Schrägstrich zur Kennzeichnung von Wortabbrüchen: *holt erstmal ne tei/ ne schiene*) werden im AutoAnnotator hinterlegt (vgl. Trevisan et al. 2014).

Abbildung 4 zeigt einen Annotationsausschnitt mit Auszeichnungen auf den Ebenen Wortabbrüche und Pausen. Der AutoAnnotator annotiert automatisch Pausen mit dem Tag *PAU* und Wortabbrüche mit dem Tag *ABB*:

	1743	1744	1745 [09:53.4]	1746	1747	1748	1749	1750 [09:55.9]
PM1 [v]								mhm okay (--)
PM1 [tok]								mhm
PM1 [pos]								NGIRR
PM1 [lemma]								mhm
PM1 [nv]								PM1 notiert „Wer
B3 [v]			dem en/ dem entsprechend kunststoffteile					
B3 [tok]	an	(--)	dem	en/	dem	entsprechend	kunststoffteile	
B3 [pos]	APPR XY		ART	ART	ART	ADJD	NN	
B3 [lemma]	an	(--)	d	d	d	entsprechend	kunststoffteile	
B3 [nv]								
Simultansprechen								
[Wortabbrueche]				ABB				
[Pausen]		PAU						

Abb. 4: Beispiel für die Annotation von Wortabbrüchen und Pausen in EXMARaLDA (1a\_05\_prozesshebung\_lagerist, Token 1743-1750)

Eine besondere Herausforderung für die automatische Annotation und Datenaufbereitung ist simultanes Sprechen. Eine automatische Identifikation der Stellen ist möglich, da in ModiKo alle Vorkommen von Simultansprechen in den Transkripten durch eckige Klammern gekennzeichnet sind. Problematisch ist allerdings die Tokenisierung dieser Stellen – wird *simultan* gesprochen, ist die in Abbildung 3 und 4 gezeigte Form der Datenaufbereitung nicht mehr möglich. Das Problem ergibt sich bei der parallelen Segmentierung in Token bei mehreren Sprechern, wie in Abbildung 5 (Token 1125-1127) dargestellt:

<sup>7</sup> Für die Implementierung, Modifikation und Umsetzung der automatischen Annotationen mit AutoAnnotator danken wir Maximilian Peiffer.

	1124	1125 [06:03.7]	1126	1127	1128 [06:04.6]	1129	1130	1131	1132	1133
PM1 [v]		stapler gewechselt werden			oder kann man mit einem stapler all					
PM1 [nv]	erhebung_Lagerist_Notiz 1 von PM1, Video.									
B3 [v]		nein			nein nein					
B3 [nv]										
Simultansprechen		PM1 & B3			PM1 & B3					
Wortabbrueche										
Pausen										

Abb. 5: Beispiel für die Annotation von Simultansprechen in EXMARaLDA (1a\_05\_prozesserhebung\_lagerist, Token 1124-1132)

Im Beispiel haben die simultanen Redeanteile von PM1 und B3 einen unterschiedlichen Tokenumfang: PM1: [*stapler gewechselt werden*] B3: [*nein*]. Der simultane Redeanteil von PM 1 (Zeile PM1[v]) umfasst drei Token, der von B3 (Zeile B3[v]) nur ein Token. Eine parallele Tokenisierung beider Zeilen unter Beibehaltung der visuellen Anzeige von Gleichzeitigkeit ist also nicht möglich. Um diese Stellen dennoch ohne Informationsverlust tokenisieren zu können, werden die identifizierten Stellen im ersten Schritt automatisch auf der Ebene *Simultansprechen* annotiert. Als Tag werden die Kürzel der simultan Sprechenden, verbunden mit einem &, vergeben (z.B. *PM1 & B3*, vgl. Abb. 5). Erst im zweiten Schritt werden simultane Äußerungen tokenisiert. Die Segmentierung in Token wird wie folgt umgesetzt (vgl. Abb. 6): Für die simultanen Redeanteile wird pro Sprecher auf der Ebene [tok] ein Fenster mit der maximal notwendigen Anzahl an Token (vorgegeben durch den längsten simultanen Redeanteil) angelegt und kürzere simultane Redeanteile an den linken Rand des Textfensters geschoben. Übrige Tokeneinteilungen (am rechten Rand des Textfensters) bleiben für die kürzeren Redeanteile leer.

	1124	1125 [06:03.7]	1126	1127	1128 [06:04.6]	1129	1130	1131	1132	1133
PM1 [v]		stapler gewechselt werden			oder kann man mit einem stapler all					
PM1 [tok]	zu	stapler	gewechselt	werden	oder	kann	man	mit	einem	stapler
PM1 [pos]	PTKA	NN	VVPP	VAINF	KON	VMFIN	PIS	APPR	ART	NN
PM1 [lemma]	zu	stapler	wechseln	werden	oder	können	man	mit	ein	stapler
PM1 [nv]	erhebung_Lagerist_Notiz 1 von PM1, Video.									
B3 [v]		nein			nein nein					
B3 [tok]		nein			nein	nein				
B3 [pos]		NGIRR			NGIRR	NGIRR				
B3 [lemma]		nein			nein	nein				
B3 [nv]										
Simultansprechen		PM1 & B3			PM1 & B3					
[Wortabbrueche]										

Abb. 6: Beispiel für Tokenisierung, Tagging und Annotation von Simultansprechen in EXMARaLDA (1a\_05\_prozesserhebung\_lagerist, Token 1124-1132)

Abbildung 6 zeigt exemplarisch die Datenaufbereitung für das Beispiel oben (Abb. 5): Der simultane Redeanteil von PM1 (Zeile PM1 [v]) umfasst drei Token, der von B3 (Zeile B3 [v]) nur ein Token. Der kürzere Redeanteil von B3 wird in der Zeile B3 [tok] an den linken Rand des Fensters geschoben (Token 1125, hervorgehoben in grün), die zwei folgenden Token (1126 und 1127) bleiben auf den Ebenen B3 [tok], B3 [pos] und B3 [lemma] leer. Der Redeanteil von PM1 (Zeile PM1 [v]) wird auf den Ebenen PM1 [pos], PM1 [lemma] und PM1 [nv] beschrieben.

ben. Das Beispiel zeigt, dass die Annotation auf der Ebene Simultansprechen *vor* der Tokenisierung geschehen muss, um die Information, *wo* simultan gesprochen wird, nicht zu verlieren.

### 5.3. Das Mehrebenenannotationsschema

Das ModiKo-Annotationsschema wurde iterativ entwickelt. Es orientiert sich an dem Mehrebenenannotationsansatz von Trevisan (2014) und adaptiert ihn gegenstandsspezifisch. Eine erste Version des Mehrebenenannotationsschemas wurde von Annotatoren (n=5) auf das Pilotkorpus angewendet. In Feedbackrunden mit Annotatoren und Experten wurde das Schema diskutiert und die Annotationen ausgewertet. Anhand der Ergebnisse wurde der Mehrebenenannotationsansatz überarbeitet. Im Annotationsschema wurden Kategorien (Ebenen und Tags) ergänzt oder Kategorien gestrichen sowie Beschreibungen und Instruktionen hinzugefügt und/oder präzisiert. Das Ergebnis wurde in Expertenworkshops validiert und dabei u.a. die Kategorie "Objekte" (weiter-)entwickelt und geschärft. Mit dem validierten Annotationsschema wurde von den Annotatoren (n=4) ein(en) Goldstandard (33.185 Token) annotiert. Ergänzend wurden ein Daten-Managementtool für den Annotationsprozess sowie Annotationsguidelines entwickelt und erprobt.

Das ModiKo-Mehrebenenannotationsschema umfasst 7 Ebenen und 25 Tags. Das vollständige Schema findet sich im Anhang dieses Beitrages. Das Schema umfasst sieben Spalten – eine Spalte links, die den Bezug zum Modell anzeigt (Modellebene) sowie die sechs Spalten des "eigentlichen" Annotationsschemas. Tabelle 2 zeigt einen exemplarischen Ausschnitt:

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang		Instruktion	Beispiel
				Token	Phrase / mehrere Token		
Wie? Manifestation	MID-Typ	SpSc_E SpSc_I	<b>Sprechen-Schreiben intrapersonell (Einzelperson)</b> Interdependenz von Sprechen und Schreiben bei einer Person  <b>Sprechen-Schreiben interpersonell</b> Interdependenz von Sprechen und Schreiben bei mehreren Personen	x	x	Annotiert werden Interdependenzen von Sprechen und Schreiben unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag „SpSc_E“ (E=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht.  Das Tag „SpSc_I“ (I=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.  B3: <u>die kleinen kugeln ( ) ich glaub um die ffffff fünf sechstausend stück pro palette (4)</u> PM1: ((PM1 <u>notiert</u> „5000 Stck. pro Palette“). Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) <u>wie viele passen auf so nen lkw</u>  PM1: <u>ah okay (5)</u> ((PM1 <u>notiert</u> „Gewichte & Gitterboxen“). Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) <u>wie viel hat eine gitterbox (-) wie viele teile sind drin</u>  B3: <u>(die werden) vom lkw sach ich mal erstmal bei uns in den warenengang</u> ((B3 zeigt links von sich ins Lager. Rekonstruktionsquelle: Video.)) <u>reingestellt (-) dann werden sie von dementsprechend sag ich mal lageristen ahm (-) entgegen genommen überprüft kontrolliert ob die ware der liefer/ ah lieferung entspricht</u> PM1: ((PM1 <u>notiert</u> „LKW entladen“ und „Lagerist überprüft“). Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))
		SpV_E SpV_I	<b>Sprechen-Visualisieren intrapersonell (Einzelperson)</b> Interdependenz von Sprechen und Visualisieren bei einer Person  <b>Sprechen-Visualisieren interpersonell</b> Interdependenz von Sprechen und Visualisieren bei mehreren Personen	x	x	Annotiert werden Interdependenzen von Sprechen und Visualisieren unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag „SpV_E“ (E=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht.  Das Tag „SpV_I“ (I=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	PM1: <u>mhmm okay (-) und dann ist das für sie (-) für ihre abteilung abgeschlossen</u> ((PM1 <u>zeichnet</u> zwei vertikale Linien hinter das Wort „Einlagern“). Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))  PM1: <u>ah okay (-) das wäre einlagern</u> ((PM1 <u>zeichnet</u> eine Trennlinie um das bisher Notierte. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))  PM1: <u>(dann) wird das teil an den maschinenbediener übergeben (8)</u> ((PM2 <u>reißt</u> ein zweites Blatt vom Block ab. PM2 <u>zeichnet</u> einen Trennungsstrich und <u>notiert</u> „Maschine übergeben“, „Einrichter“, „Maschinenbed“). Rekonstruktionsquelle: 08_Prozessmodellierung_Stufenpresse_Feinmechaniker_Notiz 2 von PM2, Video.))

Tab. 2: Ausschnitt des in ModiKo entwickelten Mehrebenenannotationsschemas

Die erste Spalte links (*Modellebene*) verortet die Annotationsebenen rechts daneben im Modell für MID-Formen (siehe Abb. 2 oben). Der Fokus richtet sich dabei auf Phänomene der Erfassung und Beschreibung von MID aus Mikroperspektive. Die makroperspektivischen Rahmungen (Situations- und Interaktionstyp) werden

im Metadatensatz der Dateien erfasst und vererben sich damit auf alle Bestandteile des Datensatzes. Die Berücksichtigung der makroperspektivischen Rahmung "Handlungsschema" ist als solche geplant, aber noch nicht im Annotationsschema umgesetzt. Es wird davon ausgegangen, dass das Handlungsschema bei der Erfassung der Funktionen von MID relevant wird. Ihre Ausarbeitung erfolgt in der nächsten Projektphase im Kontext der Bestimmung von MID-Funktionen.

Das Annotationsschema orientiert sich an den mikroperspektivischen interaktionalen Betrachtungsdimensionen. Das "Was?" (Modality-taking) wird annotiert über die Realisierungen von MID im "Wie?" (Manifestation) und "Wer?" (Akteur). Der interaktionale Effekt ("Mit wem?") ist nicht Teil des Annotationsschemas, da sich das Schema auf die Betrachtung von MID und die MID erzeugenden Akteure konzentriert.

Die zweite Spalte von links (*Ebene*) listet die in der Annotation erfassten Phänomene – basierend auf den Betrachtungsdimensionen der Modellebene. Dies ist zum einen das Zusammenspiel von Modalitäten, das im Annotationsschema als *MID-Typ* erfasst und beschrieben wird. Es sind zum zweiten Indikatoren, die auf MID *hinweisen* (MID-Indikator vorangestellt, *IV*) sowie Indikatoren, die MID *retrospektiv* verdeutlichen (MID-Indikator nachgestellt, *IN*). Sie werden aus Gründen der praktischen Handhabbarkeit in ihrer konkreten Ausprägung (sprachlich, kommunikativ, praktisch) annotiert (z.B. sprachlicher MID-Indikator vorangestellt, *SIV*). Analyserlevant sind weiter die *Objekte*, die im Kontext einer MID von den Interaktionsbeteiligten (Akteure) relevant gesetzt werden, die *Akteure* selbst sowie der *Beginn des Modality-taking*.

In der dritten Spalte (*Tag*) erscheinen die Tags (z.B. *SpSc\_I*), mit denen die in der vierten Spalte von links genannten Phänomene (*Tag-Beschreibung*) (z.B. MID-Typ *Sprechen-Schreiben interpersonell*) annotiert werden. Die in Spalte vier genannten Phänomene beziehen sich auf Kategorien der Spalte zwei (*Ebene*) in ihrer konkreten Ausprägung. MID-Typen werden beschrieben über die genutzten Ressourcen sowie die Anzahl der an der MID beteiligten Akteure (z.B. *SpSc\_E* – Sprechen-Schreiben durch Einzelperson). Die sprachlichen Indikatoren werden näher charakterisiert durch Angabe der Wortart (z.B. *SIN-V* – sprachlicher MID-Indikator nachgestellt: Verb), die kommunikativen Indikatoren werden spezifiziert durch Angabe der Art der kommunikativen Handlung (z.B. *KIV-G* – kommunikativer MID-Indikator vorangestellt: Gestik). Objekte werden auf signifikante Objekte beschränkt. Der Beginn des Modality-taking wird (als Aktion von Akteuren) spezifiziert durch die Tags *SELBST* und *FREMD* (Selbst- und Fremddinitierung).<sup>8</sup>

Die fünfte Spalte von links (*Tag-Umfang*) gibt an, welchen Umfang der mit einem Tag ausgezeichnete Ausschnitt haben kann:

- *Token*: Die Annotation umfasst lediglich eine Einheit, z.B. *malen* als sprachlichen MID-Indikator Verb.
- *Phrase/mehrere Token*: Annotiert werden analyserelevante Transkript-ausschnitte, die mehr als ein Token umfassen – etwa Transkript-ausschnitte, die Interdependenzen von Sprechen und Schreiben (*Ebene MID-Typ*) aufzeigen (Beispiel: *PM1: okay danke ((PM1 notiert "35.000-40.000 geschätzt.))* – *PM1*

<sup>8</sup> Das entwickelte Tagset berücksichtigt die Anforderungen nach Eindeutigkeit und Abgrenzbarkeit von Tags nach Leech (1993:287-280).

spricht und schreibt gleichzeitig, daher wird der gesamte Transkriptausschnitt auf der Ebene *MID-Typ* mit dem Tag *SpSc\_E* annotiert).

Die sechste Spalte (*Instruktion*) instruiert den Annotator, wie bzw. unter welchen Umständen ein Tag zu vergeben ist. Die siebte Spalte von links (*Beispiel*) verdeutlicht das bisher Gesagte (Spalte 2 bis 6) an Beispielen. Im Beispiel werden die für die Annotation relevanten Token bzw. Phrasen durch Unterstreichung markiert.

Das oben beschriebene Schema hat sich in der Handhabung als zielführend erwiesen – erste quantitative Auswertungen des Goldstandards bezogen auf die Ebenen *MID-Typ* und *MID-Akteur* zeigen eine hohe Übereinstimmung der händischen Annotationen. Dies lässt darauf schließen, dass der vorgestellte *ModiKo*-Mehrebenenannotationsansatz zur Beschreibung und Analyse von *MID* funktioniert.

## 6. Fazit und Ausblick

Der entwickelte *ModiKo*-Mehrebenenannotationsansatz erlaubt eine weitgehende Erfassung und Beschreibung von Modalitätsinterdependenzen im definierten Sinne, basierend auf der beschriebenen theoretischen Modellierung von *MID*-Formen. Das bislang entwickelte Annotationsschema fokussiert die Beschreibung der Formen von *MID* in professionellen mündlichen Interaktionssituationen bezogen auf das Fallbeispiel.

Der vorgestellte Mehrebenenannotationsansatz wird in *EXMARaLDA* umgesetzt und schließt die Lücke bisher fehlender Ansätze zur Annotation multimodaler Phänomene auf Grundlage eines multimodalen Transkripts. Der Annotationsansatz berücksichtigt dabei Besonderheiten bei der Verarbeitung von Gesprächstranskripten sowie Spontansprache (u.a. Kleinschreibung, Pausenzeichen, verkürzte Wörter, Wortabbrüche) und ermöglicht die Annotation von *MID*-Typen, *MID*-Indikatoren sowie begleitender Phänomene, wie z.B. in der Interaktionssituation genutzte Objekte. Die Annotation erfolgt in Teilen automatisiert, bei hochinterpretativen und komplexen Annotationen wie der Auszeichnung von *MID*-Typen stößt die automatische Annotation jedoch an ihre Grenzen. Ein Goldstandard (33.185 Token) wurde von vier Annotatoren annotiert und wird aktuell ausgewertet. Die Auswertung der Annotation erlaubt z.B. quantitative Aussagen über Auftreten und Verteilung von *MID*-Typen und *MID*-Indikatoren im Korpus.

Das Annotationsschema wird im Projektverlauf weiter ausgebaut werden. Bislang nicht gelöste Aufgaben betreffen u.a. die detaillierte Erfassung und Beschreibung von Objekten, die *MID* notwendig oder optional in der Interaktion begleiten. Die Signifikanz von Objekten ändert sich dynamisch in der Situation und hängt u.a. ab von der Funktionalität der Objekte für die Interaktion bezogen auf die Aufgaben der Interaktionsbeteiligten. Eine andere Herausforderung ist die Integration der Ebene Handlungsschema und ihre Ausarbeitung im Kontext der Bestimmung von *MID*-Funktionen. In diesem Kontext werden intermediäre Objekte mit zu berücksichtigen sein.

## 7. Literatur

- Beißwenger, Michael (2013): Das Dortmunder Chat-Korpus: ein annotiertes Korpus zur Sprachverwendung und sprachlichen Variation in der deutschsprachigen Chat-Kommunikation. Online-Publikation auf dem Linguistik Server Essen (LINSE). Online verfügbar:  
[http://www.linse.uni-due.de/tl\\_files/PDFs/Publikationen-Rezensionen/Chatkorpus\\_Beisswenger\\_2013.pdf](http://www.linse.uni-due.de/tl_files/PDFs/Publikationen-Rezensionen/Chatkorpus_Beisswenger_2013.pdf).
- Deutsche Forschungsgemeinschaft (DFG) (2015): Handreichung. Empfehlungen zu datentechnischen Standards und Tools bei der Erhebung von Sprachkorpora. [http://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/informationen\\_fachwissenschaften/geisteswissenschaften/standards\\_sprachkorpora.pdf](http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_sprachkorpora.pdf)
- Duncan, Susan / Rohlfing, Katharina / Loehr, Dan (2013): Multimodal Annotation Tools. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / McNeill, David (Hg.): Handbook Body-Language-Communication. Berlin / Boston: De Gruyter, 1015-1022.
- Eller, Monika / Hirschmann, Hagen (2014): Modellierung nichtstandardisierter Schriftlichkeit. In: Zeitschrift für germanistische Linguistik 42 (1), 100-110.
- Fiehler, Reinhard (1980): Kommunikation und Kooperation. Theoretische und empirische Untersuchungen zur kommunikativen Organisation kooperativer Prozesse. Berlin: Einhorn-Verlag.
- Halverson, Erica / Bass, Michelle / Woods, David (2012): The Process of Creation: A Novel Methodology for Analysing Multimodal Data. In: The Qualitative Report (17), 21.
- Hinrichs, Erhard / Bartels, Julia / Kawata, Yasuhiro / Kordoni, Valia / Telljohann, Heike (2000): The Tübingen Treebanks for Spoken German, English, and Japanese. In: Wahlster, Wolfgang (Hg.): Verbmobil: Foundations of Speech-to-Speech Translation. Berlin: Springer, 552-576.
- Jakobs, Eva-Maria / Fiehler, Reinhard / Eraßme, Denise / Kursten, Anne (2011): Industrielle Prozessmodellierung als kommunikativer Prozess. Eine Typologie zentraler Probleme. In: Gesprächsforschung 12, 223-264.
- Kipp, Michael (2014): ANVIL: The Video Annotation Research Tool. In: Durand, Jacques / Gut, Ulrike / Kristoffersen, Gjert (Hg.), Handbook of Corpus Phonology, Oxford University Press, Chapter 21, 420-436.
- Leech, Geoffrey (1993): Corpus Annotation Schemes. In: Literary & Linguistic Computing 8, 275-281.
- Neunerdt, Melanie / Trevisan, Bianka / Reyer, Michael / Mathar, Rudolf (2013a): Part-of-Speech Tagging for Social Media Texts. In: Gurevych, Iryna / Biemann, Chris / Zesch, Torsten (Hg.), Language Processing and Knowledge in the Web. Heidelberg u.a.: Springer, 139-150.
- Neunerdt, Melanie / Reyer, Michael / Mathar, Rudolf (2013b): A POS Tagger for Social Media Texts trained on Web Comments. In: Polibits (48), 61-68. [http://polibits.gelbukh.com/2013\\_48/A%20POS%20Tagger%20for%20Social%20Media%20Texts%20Trained%20on%20Web%20Comments.html](http://polibits.gelbukh.com/2013_48/A%20POS%20Tagger%20for%20Social%20Media%20Texts%20Trained%20on%20Web%20Comments.html)
- Rehbein, Ines / Schalowski, Sören (2013): STTS goes Kiez: Experiments on Annotating and Tagging Urban Youth Language. In: Journal for Language Tech-

- nology and Computational Linguistics (Themenheft "Das STTS-Tagset für Wortartentagging - Stand und Perspektiven"), 199-227.
- Rehbein, Ines / Schalowski, Sören / Wiese, Heike (2014): The KiezDeutsch Korpus (KiDKo) Release 1.0. In: Proceedings of the 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC), May 24-31, 2014. Reykjavik, Iceland.
- Reimer, Eva / Trevisan, Bianka / Eraßme, Denise / Schmidt, Thomas / Jakobs, Eva-Maria (2015): Annotating Modality Interdependencies. Proceedings of the International Conference of the German Society of Computational Linguistics and Language Technology, GSCL 2015. Universität Duisburg-Essen, Deutschland, 110-111.
- Schiller, Anne / Teufel, Simone / Stöckert, Christine / Thielen, Christine (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS. (Kleines und großes Tagset). Universität Stuttgart, Institut für maschinelle Sprachverarbeitung; Universität Tübingen, Seminar für Sprachwissenschaft.  
<http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf>
- Schmid, Helmut (1995): Improvements in Part-of-Speech Tagging with an Application to German. In: Proceedings of the ACL SIGDAT-Workshop. Dublin, Ireland.
- Schmidt, Thomas / Wörner, Kai (2009): EXMARaLDA – Creating, analyzing and sharing spoken language corpora for pragmatics research. In: Pragmatics – Quarterly Publication of the International Pragmatics Association 19.4, 565-582.
- Schneider, Wolfgang (2002): Annotationsstrukturen in Transkripten. DV-technische Strukturanforderungen für Annotate exemplifiziert an EXMARaLDA. In: Gesprächsforschung 3, 192-236.
- Schwab, Götz (2006): Transana - ein Transkriptions- und Analyseprogramm zur Verarbeitung von Videodaten am Computer. In: Gesprächsforschung 7, 70 - 78.
- Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung 10, 353-402.
- Sloetjes, Han / Wittenburg, Peter / Somasundaram, Aarthi (2011): ELAN - aspects of interoperability and functionality. In: INTERSPEECH-2011, 3249-3252.
- Telljohann, Heike / Versley, Yannick / Beck, Kathrin / Hinrichs, Erhard / Zastrow, Thomas (2013): STTS als Part-of-Speech-Tagset in Tübinger Baumbanken. In: JLCL 28, 1-16.
- Trevisan, Bianka (2014): Bewerten in Blogkommentaren. Mehrebenenannotation sprachlichen Bewertens. Dissertation. RWTH Aachen University.
- Trevisan, Bianka / Neunerdt, Melanie / Hemig, Tim / Jakobs, Eva-Maria / Mathar, Rudolf (2014): Detecting Ironic Speech Acts in Multilevel Annotated German Web Comments. In: Pre-conference workshop NLP 4 CMC: Natural Language Processing for Computer-Mediated Communication / Social Media, KONVENS 2014, 07.-10.10.2014, Hildesheim.
- Trevisan, Bianka / Reimer, Eva / Digmayer, Claas / Ullrich, Anna Valentine / Jakobs, Eva-Maria (2016): Digital Humanities und Linguistik: Herausforderungen und ihre Potenziale am Beispiel der Annotation multimodaler Daten. In: Konferenzabstracts der 3. Tagung des Verbands Digital Humanities im

- deutschsprachigen Raum e.V. (DHd 2016), 07.03.-12.03.2016, Leipzig, 382-385.
- Ule, Tylman / Hinrichs, Erhard (2004): Linguistische Annotation. In: Lobin, Henning / Lemnitzer, Lothar (Hg.), *Texttechnologie, Perspektiven und Anwendungen*. Tübingen: Stauffenburg, 217-243.
- Ullrich, Anna Valentine / Jakobs, Eva-Maria / Eraßme, Denise (2016): "ich schreib das mal hier rein ähm". Modality-taking – Schreibhinweise in professionellen mündlichen Interaktionssituationen. In: *Glottology* 7 (1): 67-88.
- Ullrich, Anna Valentine / Reimer, Eva / Jakobs, Eva-Maria (im Druck): Modalitätsinterdependenzen im professionellen Kontext: Formen und Indikatoren. Erscheint in: *ZfAL*.
- Westpfahl, Swantje (2014): STTS 2.0? Improving the Tagset for the Part-of-Speech-Tagging of German Spoken Data. In: Levin, Lori / Stede, Manfred (Hg.): *Proceedings of LAW VIII - The 8th Linguistic Annotation Workshop*. Dublin, Ireland: Association for Computational Linguistics and Dublin City University, 1-10.
- Westpfahl, Swantje / Schmidt, Thomas (2013): POS für(s) FOLK – Part of Speech Tagging des Forschungs- und Lehrkorpus Gesprochenes Deutsch. In: *Journal for Language Technology and Computational Linguistics* 1, 139-156.
- Westpfahl, Swantje / Schmidt, Thomas (2016): FOLK-Gold. – A GOLD standard for Part-of-Speech-Tagging of Spoken German. In: *Proceedings of the Tenth conference on International Language Resources and Evaluation (LREC' 2016)*. European Language Resources Association (ELRA). Portorož, Slovenia, May 23-28.
- Wiese, Heike / Freywald, Ulrike / Schalowski, Sören / Mayr, Katharina (2012): Das KiezDeutsch-Korpus. Spontansprachliche Daten Jugendlicher aus urbanen Wohngebieten. In: *Deutsche Sprache* 2, 97-123.
- Wittenburg, Peter (2008): Preprocessing multimodal corpora. In: Lüdeling, Anke / Kytö, Merja (Hg.), *Corpus Linguistics. An international handbook* (1), Berlin: De Gruyter, 664-685.

Prof. Eva-Maria Jakobs  
Textlinguistik und Technikkommunikation  
Institut für Sprach- und Kommunikationswissenschaft  
RWTH Aachen  
Campus-Boulevard 57  
52074 Aachen

[e.m.jakobs@tk.rwth-aachen.de](mailto:e.m.jakobs@tk.rwth-aachen.de)

Anhang: ModiKo-Mehrebenenannotatiosschema

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
				Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
Wie? Manifestation	MID-Typ	SpSc_E SpSc_I	<b>Sprechen-Schreiben intrapersonell (Einzelperson)</b> Interdependenz von Sprechen und Schreiben bei einer Person  <b>Sprechen-Schreiben interpersonell</b> Interdependenz von Sprechen und Schreiben bei mehreren Personen	Token  x	Annotiert werden Interdependenzen von Sprechen und Schreiben unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag "SpSc_E" (=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht. Das Tag "SpSc_I" (=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	B3: "die kleinen Kugeln ( ) ich glaub um die ffffff fünf sechstausend Stück pro palette (A)" PM1: ((PM1 notiert „5000 Stck. pro Palette“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) "wie viele passen auf so nen Ikw" PM1: "ah okay (S)" (( PM1 notiert „Gewichte & Gitterboxen“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) "wie viel hat eine Gitterbox (-) wie viele teile sind drin" B3: "(die werden) vom Ikw sach ich mal erstmal bei uns in den warenengang" ((B3 zeigt links von sich ins Lager. Rekonstruktionsquelle: Video.)) "reingestellt (-) dann werden sie von dementsprechend sag ich mal Lageristen ähm (-) entgegen genommen überprüft kontrolliert ob die ware der liefer/ äh lieferung entspricht" PM1: ((PM1 notiert „LKW entladen“ und „Lagerist überprüft“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))
Wie? Manifestation	MID-Typ	SpV_E SpV_I	<b>Sprechen-Visualisieren intrapersonell (Einzelperson)</b> Interdependenz von Sprechen und Visualisieren bei einer Person  <b>Sprechen-Visualisieren interpersonell</b> Interdependenz von Sprechen und Visualisieren bei mehreren Personen	x  x	Annotiert werden Interdependenzen von Sprechen und Visualisieren unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag "SpV_E" (=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht. Das Tag "SpV_I" (=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	PM1: "mhm okay ( ) und dann ist das für sie ( ) für ihre abteilung abgeschlossen" (PM1 zeichnet zwei vertikale Linien hinter das Wort „Einlagern“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))  PM1: "ah okay (-) das wäre einlagern" ((PM1 zeichnet eine Trennlinie um das bisher Notierte. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))  PM1: "(dann) wird das teil an den maschinenbediener übergeben (8)" (PM2 reißt ein zweites Blatt vom Block ab. PM2 zeichnet einen Trennungsstrich und notiert „Maschine übergeben“, „Einrichter“, „Maschinenbed“. Rekonstruktionsquelle: 08_Prozessmodellierung_Stufenpresse_Feinmechaniker_Notiz 2 von PM2, Video.))
Wie? Manifestation	MID-Typ	VSc_E VSc_I	<b>Visualisieren-Schreiben intrapersonell (Einzelperson)</b> Interdependenz von Schreiben und Visualisieren bei einer Person  <b>Visualisieren-Schreiben interpersonell</b> Interdependenz von Schreiben und Visualisieren bei mehreren Personen	x  x	Annotiert werden Interdependenzen von Schreiben und Visualisieren unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag "VSc_E" (=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht. Das Tag "VSc_I" (=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	B3: "und der zweite schritt ist dann halt/ dann wird die ware sach ich mal ins system ( ) buchungstechnisch eingebucht (-)" PM1: "ach so moment (S)" ((PM1 streicht auf dem Blatt Notiertes: „Einbuchen im System“ durch, und notiert an anderer Stelle auf dem Blatt „System erfasst“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.)) PM1: "so erfasst und dann wird eingelagert ähm mit nem stapler ähm muss da irgendwie noch von stapler zu (stapler gewechselt werden)" ((PM1 notiert „Einlagern“ und unterstreicht das Wort. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.)) B6: "das/ brauchen die doppelt zeit weil sie länger gefährt werden müssen ( ) aber normalerweise ( ) variiert höchstens die eloakzeit oder die glanzzeit je nachdem welches teil es ist" ((PM1 notiert „Beizen“ und streicht es dann wieder durch. Rekonstruktionsquelle: 11_Prozesserhebung_Eloxieren_Galvaniseur_Notiz 1 von PM1, Video.))

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
				Phrase / mehrere Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
Wie? Manifestation	MID-Typ	SpScv_E SpScv_I	<b>Sprechen-Schreiben-Visualisieren</b> intrapersonell (Einzelperson) Interdependenz von Sprechen, Schreiben und Visualisieren bei einer Person  <b>Sprechen-Schreiben-Visualisieren</b> interpersonell Interdependenz von Sprechen, Schreiben und Visualisieren bei mehreren Personen	Token  x	Annotiert werden Interdependenzen der drei Modalitäten unabhängig von der Reihenfolge ihres sequentiellen oder simultanen Auftretens.  Das Tag „SpScv_E“ (E=Einzelperson) wird vergeben, wenn sich die MID auf eine Person bezieht. Das Tag „SpScv_I“ (I=interpersonell) wird vergeben, wenn sich die MID zwischen mehreren Personen ereignet.	PM1: „ja (4) (L...) hier würd ich/ hier würd ich noch oben (-) auch info an fahrer ne“ (PM1 zeichnet ein Rechteck auf die Verbindungslinie zwischen „Liefererschein korrekt Nein“ und „LKW beladen“, zieht eine Linie in die Swirlane „LKW-Fahre“ und schreibt hier „Info an Fahrer geben“; Rekonstruktionsquelle: 11_Prozessmodellierung_Ubergabe an PM4_Notiz von PM1 (vgl. Objekt 6) Video.) PM1: ((PM1 streicht Pfeil durch und notiert „Lagerist“; Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) "ich schreib dann mal lagerist damit [ichs (?)]" PM1: "so dann wenns erfasst ist dann nimmt man die zettel oder was bekommt der lagerist dass er irgendwie/" B3: "der kriegt einen transportauftrag" (B3 zeichnet mit den Händen ein Rechteck in die Luft; Rekonstruktionsquelle: Video) PM1: "mhmm" ((PM1 notiert „Transportauftrag“; Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.)) B3: "so genannten wo die lagerplätze mit dementsprechenden lagereinheiten sich drauf befinden (-) und dementsprechend klebt er diese lagereinheiten an die äh vorgesehene paletten" PM2: "machten wer mal informieren (-) okay und dann" ((PM2 notiert „Schaden dokumentieren informieren“; Rekonstruktionsquelle: 06-6a_Prozessmodellierung_Lagerist_Notiz 1 von PM2, Video.)) PM1: "abhängen und dann das luftrocknen das is (-) ich mal das mal so in klammern dass dann zusammen luft (-) trock (-) nen (-) okay so ((PM1 notiert „Abhängen Luftrocknen“; Rekonstruktionsquelle: 02-2a_Prozessmodellierung_Meister_Notiz 1 von PM1, Video.))
Wie? Manifestation	MID-Indikator sprachlich	SIV-V	<b>Sprachlicher MID-Indikator vorangestellt:</b> Verb Verb, das auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM1: ((PM1 streicht Pfeil durch und notiert „Lagerist“; Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.)) "ich schreib dann mal lagerist damit [ichs (?)]" PM1: "I...I" Ich hab eben ((PM1 klebt einen Post-it auf das Dokument 03_Prozesserhebung_Planer_Notiz von Planer und notiert „Tiefziehen“; Rekonstruktionsquelle: Video.)) schon gelernt dass die kugel tiefgezogen wird" B2: "ja genau ((PM1 und B2 lachen; Rekonstruktionsquelle: Video.))" PM1: "das schreib ich mir als beispiel schon mal hier hin (-) tiefziehen (-) und dass sie die aktivitäten dann so verbinden (-) und vielleicht ähm (-) auch mal (-) ganz notwendig die losgrößen hinschreiben zum ((PM1 klebt ein weiteres Post-it auf das Blatt vor sich und zeichnet einen Pfeil zwischen den beiden Post-its. Rekonstruktionsquelle: Video.)) beispiel vierhundert stück (-) wenn sies wissen" B2: "mhmm"
Wie? Manifestation	MID-Indikator sprachlich	SIV-ADV	<b>Sprachlicher MID-Indikator vorangestellt: Adverb</b> Adverb, das auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM1: "ähm (-) ja sie die aktivitäten (-) gleich benennen die notwendig sind/ ich schreib das mal hier rein ähm (-) ich hab eben" ((PM1 klebt einen Post-it auf das Dokument 03_Prozesserhebung_Planer_Notiz von Planer und notiert „Tiefziehen“; Rekonstruktionsquelle: Video.)) "schon gelernt dass die kugel tiefgezogen wird" PM1: "mhmm" ((PM1 ist das für sie (-) für ihre abteilung abgeschlossen" ((PM1 zeichnet zwei vertikale Linien hinter das Wort „Einlagern“; Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.))

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
Wie? Manifestation				Phrasen / mehrere Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
Wie? Manifestation	MID-Indikator sprachlich	<b>SIN-ADV</b>	<b>Sprachlicher MID-Indikator nachgestellt:</b> <b>Adverb</b> , das auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	PM1: "ne ne das geht auch nicht also weil das wär ja ne andere aktivität ähm das würd ich noch mal einzeln hier wieder einzelne aktivitäten (PM1 zeichnet verschiedene Striche und Halbkreise in das Dokument 02_Prozessmodellierung_Übergabe an PM4_Notiz von PM1 (vgl. Objekt 5). Rekonstruktionsquelle: 02_Prozessmodellierung_Übergabe an PM4_Notiz von PM1, Video.)" PM4: "nhm" PM1: "weil das sind ja andere bauteile hier [würd]" PM4: "okay ja" PM1: "ich auch wirklich sagen was es für bauteile sind" PM4: "ja" B2: "kunststoff (-) gut hier (-) granulat (5) (B2 notiert „Granulat“. Rekonstruktionsquelle: 03_Prozesserhebung_Planer_Notiz von Planer, Video.) und dann kommen noch bei (-) bei der variante m/ bei den kugelbeteiligten" (B2 notiert „Kugelbeteiligte“. Rekonstruktionsquelle: 03_Prozesserhebung_Planer_Notiz von Planer, Video.)
Wie? Manifestation	MID-Indikator sprachlich	<b>SIV-SUB</b>	<b>Sprachlicher MID-Indikator vorangestellt: Substantiv</b> Substantiv, das auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM2: "also wäre überprüfen" PM1: "wäre überprüfen (-) a/ anhand liefertschein (5)" PM2: ((PM2 notiert „Wäre überprüfen“, darunter „Lieferschein ok“. Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 1 von PM2, Video.) B3: "und der zweite schritt ist dann halt/ dann wird die ware sach ich mal ins system (-) buchungstechnisch eingebucht (-)" PM1: "ach so moment (5)" ((PM1 streicht auf dem Blatt Notiertes: „Einbuchen im System“ durch und notiert an anderer Stelle auf dem Blatt „System erfasst“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.)
Wie? Manifestation	MID-Indikator sprachlich	<b>SIN-SUB</b>	<b>Sprachlicher MID-Indikator nachgestellt: Substantiv</b> Substantiv, das auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	PM1: ((PM1 streicht Pfeil durch und notiert „Lagerist“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 2 von PM1, Video.) "ich schreib dann mal lagerist damit (chs (?))" B1: "ok" ((Nimmt den Stift in die Hand. B1 notiert „Alu“. Rekonstruktionsquelle: 01_Prozesserhebung_Meister_Notiz 1 von PM1, Video.) PM1: "genau" B1: "also als erstes isses ganz klar das alu"
Wie? Manifestation	MID-Indikator sprachlich	<b>SIV-P</b>	<b>Sprachlicher MID-Indikator vorangestellt: Partikel</b> Partikel, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM1: "nhm okay (-) und dann ist das für sie (-) für ihre abteilung abgeschlossen" ((PM1 zeichnet zwei vertikale Linien hinter das Wort „Einlagern“. Rekonstruktionsquelle: 05_Prozesserhebung_Lagerist_Notiz 1 von PM1, Video.) PM1: "äh ok gut (-) und dann sieht das so aus" ((PM1 dreht das Blatt zu sich und zeichnet ein Rechteck in das er „Teile anlegen“ schreibt und daneben einen Pfeil zeichnet. Rekonstruktionsquelle: 01_Prozesserhebung_Meister_Notiz 1 von PM1, Video.)

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
Wie? Manifestation	MID-Indikator sprachlich	SIN-P	Sprachlicher MID-Indikator nachgestellt: Partikel, die auf eine MID im Umfeld hinweist.	Token x Phrasen / mehrere Token	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.  B5: "und dann werden die da direkt aufgeclipst" PM1: "ah okay" (PM1 notiert „an Anlage clippen“. Rekonstruktionsquelle: 09_Prozesserhebung_AufsteckenAbhängen_Feinnmechaniker_Notiz von PM1, Video.) B5: "macht er jetzt grad da" PM1: "mhmm" B5: "die kugeln (-)" PM1: "äh sehr gut (-)" PM1: "mhmm dann gehts in die kunststoffabteilung" PM2: "okay" ----- PM2: "mhmm (-) qualitätskontrolle" (PM2 notiert „OK“. Rekonstruktionsquelle: 04_Prozessmodellierung_Planer_Notiz von PM2, Video.) PM1: "mhmm dann gehts in die kunststoffabteilung" PM2: "okay"
Wie? Manifestation	MID-Indikator kommunikativ	KIV-B	Kommunikativer MID-Indikator vorangestellt: Blick Blickbewegung eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM1: "also er macht es so er clipst eins und clipst dann automatisch wieder eins drauf" B6: "ja (-) genau (-) es wird dann quasi (-) einfach (-) so (-)" ((PM1 und B6 stehen auf und schauen durch das Fenster in die Produktionshalle. Rekonstruktionsquelle: Video)) ((PM1 notiert „1) Teile fahren“, „2) ent!“; Rekonstruktionsquelle: 11_Prozesserhebung_Eloxieren_Galvaniseur_Notiz 1 von PM1, Video.)) ----- PM2: "das gleiche gilt/ genau das gilt auch für (-) dann brauchen wir nicht mal modellieren für ähm pumpen äh diesen für diesen gilt das gleiche und für gewichte ist das der gleiche vorgang können wir noch mit draufschreiben (5)" ((PM1 schaut in seine Unterlagen. PM2 schaut in das Dokument 06+6a_Prozessmodellierung_Lagerist_Notiz 4 von PM2. Rekonstruktionsquelle: Video.)) "diesen und [gewichte]" PM1: "mhmm diesen und gewichte (6)" PM2: "mhmm (-)" ((PM2 notiert „Düsen/ Gewichte“. Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 4 von PM2, Video.))

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
Wie? Manifestation				Token Phrasen / mehrere Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
Wie? Manifestation	MID-Indikator kommunikativ	KIN-B	<b>Kommunikativer MID-Indikator nachgestellt: Blick</b> Blickbewegung eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	PM2: "okay" (PM2 notiert „Problem beheben (dann 1)“, „1“: Rekonstruktionsquelle: 10_Prozessmodellierung_AufsteckenAbhängen_Feinmechaniker_Notiz 2 von PM2, Video.) PM1: <u>„genau dann haben wir noch hier bei der mitarbeiterin/ wo bringt sie das teil weg (-) wo fährt sie weg (-) äh“</u> (PM1 schaut auf das Blatt vor PM2. Rekonstruktionsquelle: Video.) PM2: „da“ PM1: „genau davor kommt noch label (-) wird ähm auf den wagen gelecht das label wird hier aus der box geholt und auf den wagen gelecht (-)“ ----- PM1: „das werk (-) jetzt hatten wir noch zwei sachen was interessant is jetzt hier noch mal (-) vielleicht so kenndaten (-)“ (PM1 notiert „10 000 Stk./Schicht“, Rekonstruktionsquelle: 02+2a_Prozessmodellierung_Meister_Notiz 1 von PM1, Video.) „also er schafft“ (PM1 schaut auf das Blatt vor PM2. Rekonstruktionsquelle: Video.) „zehntausend [stück]“ PM2: „[nhm]“ PM1: „pro schicht bei der grauen kugel“ (PM1 notiert „10 000 Stk./Schicht“, Rekonstruktionsquelle: 02+2a_Prozessmodellierung_Meister_Notiz 1 von PM1, Video.)
Wie? Manifestation	MID-Indikator kommunikativ	KIV-G	<b>Kommunikativer MID-Indikator vorangestellt: Gestik</b> Gestik eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	B1: <u>„[meine persönlichen]“ (B1 zeigt auf sich und zieht das vor ihm liegende Blatt näher zu sich heran. Rekonstruktionsquelle: Video.)</u> PM1: <u>„ne von der abteilung an sich was in ihrer abteilung passiert (-) damit die kugel (-) produziert werden muss“ (PM1 zeigt mit der ganzen Hand auf den Tisch. Rekonstruktionsquelle: Video.)</u> B1: <u>„ok“ (Nimmt den Stift in die Hand. B1 notiert „Alu“ Rekonstruktionsquelle: 01_Prozesserhebung_Meister_Notiz 1 von PM1, Video.)</u> ----- B9: <u>„und sich dann anschauen“ (B9 zeigt auf ein Memoboard. Rekonstruktionsquelle: Video.)</u> „und dann ähm ja hamms=wers dann hier (-) (?) macht se dann ähm für vierzig teile sechzig sekunden ne fünfundvierzig da stehts“ PM1: <u>„nhm okay“ (PM2 notiert „45 Sekunden zu Teile“, Rekonstruktionsquelle: 18_Prozesserhebung_Montage_Notiz 2 von PM2, Video.)</u>

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
Wie? Manifestation				Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
	MID-Indikator kommunikativ	KIN-G	<b>Kommunikativer MID-Indikator</b> nachgestellt: Gestik Gestik eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x Phrase/ mehrere Token	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	B1: "montage (-)" ((B1 notiert „Montage“; Rekonstruktionsquelle: O1_Prozesserhebung_Meister_Notiz 1 von PM1, Video.) und da spielen verschiedene Komponenten mit äh en ganz großen (-) die spielen die mit ein (-) also einmal die düsen" ((B1 notiert „Düsen“; Rekonstruktionsquelle: O1_Prozesserhebung_Meister_Notiz 1 von PM1, Video.) PM1: "nhm" B1: "düsen" PM1: <b>was heißt düsen in dem fall</b> ((PM1 zeigt auf das geschriebene Wort „Düse“; Rekonstruktionsquelle: Video.) B1: "düsen (-) das sind halt bei der oberschale (-) ist die düse drin wo das [parfüm rausgesprüht wird] (-)" ----- PM1: " das wäre halt schon für mich ne hilfe für den prozess wenn ich ihn mir unten anschau" ((B1 zeichnet einen Pfeil hinter die Rechtecke und schreibt das Wort „Aufstecken“; Rekonstruktionsquelle: O1_Prozesserhebung_Meister_Notiz 1 von PM1, Video.) B1: " aufstecken (-) ja" PM1: aufstecken heißt in die maschine aufstecken [das is/]" ((PM1 zeigt auf das Wort Aufstecken und gestikuliert mit den Händen. Rekonstruktionsquelle: Video.)
Wie? Manifestation	MID-Indikator praktisch	PIV	<b>Praktischer MID-Indikator vorangestellt</b> Praktische Handlung eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn eine MID folgt bzw. sie im Kontext einer MID stehen.	PM1: "was wir eben gesehen haben auch (5) genau richtig lkw entladen die sendung wird komplett entladen und auf-em zwischenlagerplatz abgestellt lkw entladen absch/ schreib einfach abstellen auf zwischenlagerplatz genau irgendwie so was (9)" PM2: "ach" ((PM2 klebt einen Zettel auf das Blatt vor sich; Rekonstruktionsquelle: Video.) PM1: "lagerist genau" PM2: "warte" ((PM2 notiert „auf Zwischenlager abstellen“; Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 1 von PM2, Video.) ----- PM1: "ne von der abteilung an sich was in ihrer abteilung passiert (-) damit die kügel (-) produziert werden muss" ((PM1 zeigt mit der ganzen Hand auf den Tisch. Rekonstruktionsquelle: Video.) B1: " ok" ((Nimmt den Stift in die Hand; B1 notiert „Alu“; Rekonstruktionsquelle: O1_Prozesserhebung_Meister_Notiz 1 von PM1, Video.) PM1: "genau" B1: "also als erstes isses ganz klar das alu"

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
				Token Phrasen/ mehrere Token		<i>Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.</i>
Wie? Manifestation	MID-Indikator praktisch	<b>PIN</b>	<b>Praktischer MID-Indikator nachgestellt</b> Praktische Handlung eines Interaktionsbeteiligten, die auf eine MID im Umfeld hinweist.	x	MID-Indikatoren werden nur annotiert, wenn sie auf eine MID folgen bzw. sie im Kontext einer MID stehen.	PM1: " <u>die/äh</u> ( <u>) also hier</u> (PM1 zeichnet den Ansatz eines Rechteckes und zeigt darauf. Rekonstruktionsquelle: <i>O1_Prozesseriebung_Meister_Notiz 1 von PM1, Video</i> .) <u>findet ne Qualitätskontrolle stat</u> " B1: ((B1 zeigt auf das Blatt und nimmt den Stift; Rekonstruktionsquelle: Video.)) "genau das sind wie gesacht an/ das ist ein zwp bereich da [werden]" ----- PM1: " <u>ah ok gut</u> ( <u>) ist das was hier stattfindet passiert das in der einen maschine die hier unten/</u> " ((B1 malt mit dem Finger einen Kreis auf die Zeichnung. Rekonstruktionsquelle: Video.)) B1: " <u>ja</u> ( <u>)"PM1: "<u>nhim ok</u> (<u>) [also]" B1: "<u>also</u> wir ham für die wir ham ein wir haben zwei [maschinen]" PM1: "<u>nhim</u>" B1: "<u>xkugel einmal oberteil unterteil oberteil bekommen die düse</u> (<u>)"((B1 nimmt den Stift auf und zeigt auf die Zeichnung. Rekonstruktionsquelle: Video.))</u></u></u>
Wie? Manifestation	Objekt	<b>SIG-Obj</b>	<b>Signifikantes Objekt</b> Mindestens von einem Interaktionsbeteiligten genutzter physischer Gegenstand mit Relevanz für den folgenden Gesprächsverlauf.	x	Fall 1: Ist das Objekt Thema in der Interaktion (v-Ebene), wird das Objekt tokengenau annotiert. ----- Fall 2: Wird das Objekt in der Beschreibung kommunikativer und praktischer Handlungen (nv-Ebene) genannt, wird die gesamte Phrase der Beschreibung annotiert.	PM1: " <u>düsen montieren</u> (ID) gewicht einsetzen ( <u>) so (<u>) okay dann betrachten wer jetzt im folgenden nur seinen bereich das andere machen wer ja dann getrennt (<u>) nehmen wer mal den stift hier (<u>) so</u>" (PM1 nimmt einen stift vom tisch. Rekonstruktionsquelle: Video.)) B3: "<u>genau richtig genau</u> (?)" (PM1 nimmt ein Klemmbrett vom Boden auf. Rekonstruktionsquelle: Video.))</u></u></u>
Wer? Akteur	Akteur	(Kürzel des Akteurs/ der Akteure)	<b>Akteur(e) der MID</b> Ein oder mehrere Interaktionsbeteiligte, die Sprechen, Schreiben und/oder Visualisieren kombinieren (Entstehung einer MID).	x	Annotiert wird jeweils der gesamte Umfang des dazu gehörigen MID-Typs. Als Tag werden die Kürzel der Interaktionsbeteiligten vergeben.	B2: " <u>bereich elox</u> ( <u>) dann werden die äh kugeln (<u>) kugeln dann/ müssen wieder abgehangen werden"</u> (B2 reißt einen Post-it ab und klebt ihn auf das Blatt vor sich. Rekonstruktionsquelle: Video.))</u>

Modell-ebene	Ebene	Tag	Tag-Beschreibung	Tag-Umfang	Instruktion	Beispiel
Wert? Akteur	Beginn des Modality-taking	SELBST	Selbstinitiiierung Ein Interaktionsbeteiligter macht die Nutzung der von ihm gewählten Ausdrucksressource sprachlich explizit.	Token	Annotiert eine Selbstinitiiierung auch dann, wenn keine MID folgt, denn Initiierungen können in der Interaktionssituation auch folgenlos bleiben.	Für die Vergabe des Tags relevante Token und/oder Phrasen sind in den Beispielen unterstrichen.  PM1: „ja das muss ich mir mal aufschreiben“ ((PM1 notiert „Kopf wechseln“). Rekonstruktionsquelle: 09_Prozesserhebung_AufsteckenAbhängen_Feinnmechaniker_Notiz von PM1, Video))  PM2: „[da müssten] wir eigentlich hier ne [raute machen]“ PM1: „[ja genau]“ PM2: „[kw leer]“ ((PM2 skizziert eine Raute und Iteration (s. K3-Userguide) und notiert „LKW leer“). Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 1 von PM2, Video.))  PM2: „dann würd ich das mal hier hinsetzen!“ ((PM2 notiert „LKW beladen“). Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 1 von PM2, Video.))
				Phrase / mehrere Token		
Wert? Akteur	Beginn des Modality-taking	FREMD	Fremdinitiiierung Ein Interaktionsbeteiligter fordert einen anderen Interaktionsteilnehmer explizit auf, eine Ausdrucksressource zu nutzen.	Token	Annotiert wird eine Fremdinitiiierung auch dann, wenn keine MID folgt, denn Initiierungen können in der Interaktionssituation auch folgenlos bleiben.	PM1: „ja (→) wieder synchrone Kommunikation (15)“ ((PM2 klebt einen Zettel auf das Blatt vor sich. Rekonstruktionsquelle: Video.)) „schreib einfach (L) Schaden/ Handhabung des Schadens oder wie auch immer (16) genau (5)“ PM2: ((PM2 notiert „Schaden handhaben“). Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 2 von PM2, Video.))  PM1: „ne demach kannste sagen/ hier machste noch so ne entscheidung (.)“ zwischenlagerplatz leer ja nein wenn nein gehts wieder hier hoch (27)“ PM2: „okay (9)“ ((PM2 notiert „Zwischenlagerplatz leer?“). „ja“, „nein“). Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 3 von PM2, Video.))  PM1: „gut das wäre das dann sind die prozesse soweit fertig dann haben wir nur noch hier die mengen (.) die ich hier noch aufgeschrieben habe das wir die vielleicht da unten noch hinschreiben (3) diesen sind zehntausend stück (-) je karton (11) fünfzigtausend stück je palette (5) fünfzig bis hunderttausend stück je palette tschuldigung (9) genau (-) auf eine palette passt ein aluband also ein coil (8) kann montags dienstags mittwochs angeliefert werden (8)“ ((PM2 notiert „Düsen: 10.000 Stück/Karton, 50.000 bis 100.000 Stück/Palette, 1 Coil Alu pro Palette (Mo, Di, Mi)“). Rekonstruktionsquelle: 06+6a_Prozessmodellierung_Lagerist_Notiz 8 von PM2, Video.))
				Phrase / mehrere Token		

Veröffentlicht am 26.7.2017

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.