- Pausensituation
- Essenssituation

Die Berücksichtigung von wechselnden Situationen ermöglichte, dass die Kinder mit verschiedenen Interaktionspartner:innen (Erwachsene:r-Kind-Interaktion, Kind-Kind-Interaktion) in Kontakt kommen, sodass der natürliche Sprachgebrauch abgebildet werden konnte. Darüber hinaus orientierte sich die Datenerhebung auch an für das Kind bedeutsamen Situationen, die außerhalb des Unterrichts stattfanden (Heidtmann, 1988; Jeuk, 2003, S. 126). Darum wurde auf Videografien verzichtet. Videografien hätten eine Begrenzung auf einen ausgewählten kommunikativen Kontext (z.B. Unterricht) zur Folge, wodurch die Vielfältigkeit kommunikativer Situationen nicht erfasst werden könnte. Zudem wurde bewusst der Fokus auf lautsprachliche Äußerungen der Kinder gelegt, wenngleich dadurch die Kontextfaktoren nicht umfänglich abgebildet werden konnten.

Die erhobenen Daten wurden nach der Datenschutz-Grundverordnung (EU-DSGVO, 2016) archiviert, verwaltet und weiterverarbeitet.

11.3 Datenaufbereitung

Die gewonnenen Sprachproben wurden über ein einheitliches Transkriptionssystem (Anhang C) mit 14 Modulen durch studentische Projektmitarbeiter:innen transkribiert, welches in Anlehnung an Fuß und Karbach (2014) sowie Dresing und Pehl (2015) unter Berücksichtigung von Transkriptionsregeln aus der qualitativen Sozialforschung entwickelt wurde. Darüber hinaus wurden Transkriptionsverfahren aus deutschsprachigen und internationalen Kernvokabularstudien eingearbeitet (Boenisch, 2014b; Robillard et al., 2014; Trembath et al., 2007). Ausgeschlossen wurden Wörter, wie *Äh*, *Oh* usw. (auch Shin & Hill, 2016, S. 420). Eine leichte bis vollständige Sprachglättung wurde bei umgangssprachlichen Ausdrucksweisen (z.B. »mal« anstatt »ma«) sowie offensichtlich nicht-normgerechten Äußerungen auf der Wortebene angestrebt (z.B. ein Kind sagt *auf* anstelle von *auch*, aus dem Kontext ist aber ersichtlich, dass die unterschiedlichen Bedeutungen bekannt sind, allerdings der »ch«-Laut nicht gebildet werden kann). Ein fehlerhafter Satzbau wurde nicht geglättet. Die Umgebungssprache wurde nicht transkribiert. Vor diesem Hintergrund konnte auf komplexere Transkriptionssysteme aus der Gesprächsforschung (z.B. HIAT; Rehbein, Schmidt, Meyer, Watzke & Herkenrath, 2004) verzichtet werden (Mempel & Mehlhorn, 2014, S. 147ff.; Ziem & Lasch, 2013, S. 72). Die Transkription der Sprachaufnahmen erfolgte in MAXQDA Standard (VERBI Software, 2017). Pro Untersuchungsteilnehmer:in wurde ein Transkript (txt-Datei) erstellt. MAXQDA Standard wurde als besonders geeignet für die Datenaufbereitung eingestuft, weil sich Datenaufbereitung und -auswertung innerhalb eines Programms kombinieren lassen. Über das Erweiterungsprogramm MAX Dictio ließ sich die quantitative Analyse der genutzten Wörter und Wortkombinationen durchführen.

11.4 Datenauswertung

Für die Datenauswertung wurden wissenschaftliche Standards aus der quantitativen Korpusanalyse berücksichtigt (u.a. Mezger et al., 2016; Perkuhn et al., 2012). Das Zählen von Einheiten ist die Grundlage quantitativer korpusanalytischer Methoden (Meißner et al., 2016, S. 309f.). Die Abläufe in der Datenauswertung waren daher mit dem Ziel verbunden, die Transkripte so aufzubereiten, dass Listen über häufige Lemma-Types (Nennform des Lexems, z.B. Nominativ Singular bei Substantiven) und Wortkombinationen generiert werden konnten. Anhand der Listen konnten verschiedene formalsprachliche Aspekte (u.a. frequente und weniger frequente Einheiten, Meißner et al., 2016, S. 310) des natürlichen mündlichen Sprachgebrauchs gewonnen werden (Tab. 32).

Tab. 32: formalsprachliche Aspekte der Korpusanalyse

Bezeichnung	**Kernvokabular**
Token	Anzahl der insgesamt verwendeten Wörter (Lexeme)
Lemma-Types	Anzahl der verschiedenen Wörter (Lexeme)
Häufigkeit	Absolute Häufigkeit pro Wort
Rang	Wörter werden im Hinblick auf die Häufigkeit absteigend sortiert (größter Wert = Rang 1).
Streuung (S)	Anzahl aller Fälle pro Lemma-Types
Bezeichnung	**feste Wortkombinationen**
Token	Anzahl der insgesamt verwendeten Wortkombinationen
Types-Kombinationen	Anzahl der verschiedenen Wortkombinationen
Häufigkeit	Absolute Häufigkeit pro Wortkombination
Rang	Wortkombinationen werden im Hinblick auf die Häufigkeit absteigend sortiert (größter Wert = Rang 1).
Streuung (S)	Anzahl aller Fälle pro Types-Kombination

Im Folgenden werden die Schritte der Datenauswertung für die Analyse des Kernvokabulars sowie für die festen Wortkombinationen beschrieben.

Kernvokabular

Das Auswertungsschema für die Analyse des Kernvokabulars war stark angelehnt an die Untersuchung von Boenisch (2014b), um die gewonnenen Ergebnisse mit dem Referenzkorpus aus Boenisch (2013, Kl. 2 und 4) vergleichen zu können (Kap. 10.1). Abweichend zum Analyseschema von Boenisch (2014b) wurde in der vorliegenden Arbeit *machen* nicht zu den Hilfsverben gezählt, sondern als Vollverb behandelt (S. 14ff.). Ebenso wurde *werden* als Hilfsverb gezählt.