

Kodieren und Auswerten von Kategorien in der Interaktionalen Linguistik

Nadine Proske 

Eingegangen: 31. Oktober 2023 / Angenommen: 30. Dezember 2023
© The Author(s) 2024

Zusammenfassung Der vorliegende Artikel thematisiert das Zusammenspiel von qualitativen, interpretativen Praktiken und quantifizierenden Methoden in der Interaktionalen Linguistik. Dieser Forschungsansatz arbeitet mit Daten aus gesprochenen Interaktionen, die in Korpora aufbereitet sind. Die Auswertung der Korpusdaten umfasst sowohl Schritte, die aus der Analyse von Schriftkorpora bekannt sind, als auch solche Schritte, die spezifisch für die Analyse mündlicher Daten sind. Ein zentraler Prozess, der sowohl bereits interpretativ ist als auch der eigentlichen Ergebnisinterpretation vorausgeht, ist die Kodierung der Korpusbelege. Kodiert werden in der Interaktionalen Linguistik sowohl klassische sprachwissenschaftliche, z. B. syntaktische Kategorien als auch für Ansätze zur gesprochenen Interaktion spezifische Kategorien, z. B. im Bereich der Prosodie und der verbalen Handlungen. Der Artikel stellt zunächst vorhandene Korpora, typische Fragestellungen und methodische Vorgehensweisen der Interaktionalen Linguistik vor. Anschließend wird anhand zweier Beispiele aus eigenen interaktionslinguistischen Untersuchungen zur verbalen Pseudo-Koordination im Deutschen (*Sie können sich nicht hinstellen und sagen: »Das geht nicht.«*) gezeigt, welche interpretativen Herausforderungen Kodierkategorien und Quantifizierungen im Bereich der Prosodie und Pragmatik bergen. Abschließend wird reflektiert, inwiefern der Begriff der Hermeneutik für die diskutierten interpretativen Prozesse passend ist.

Schlüsselwörter Interaktionale Linguistik · Kategorisieren · Kodieren · Korpus · Pseudo-Koordination · Quantifikation

✉ Nadine Proske

Abteilung Pragmatik, Leibniz-Institut für Deutsche Sprache Mannheim, Mannheim, Deutschland
E-Mail: proske@ids-mannheim.de

Coding and Interpreting Categories in Interactional Linguistics

Abstract This paper addresses the interplay of qualitative, interpretive practices and quantitative methods in Interactional Linguistics. This linguistic approach uses data from corpora containing spoken interactions. The analysis of the corpus data comprises steps known from the work with written corpora as well as steps that are specific to the analysis of spoken data. One central process, which is interpretive, but also precedes the interpretation of the final results, is the coding of instances of a phenomenon. Coding in Interactional Linguistics includes classical categories from linguistics, for example syntactic ones, as well as categories that are specific to approaches investigating spoken interactions, for example prosodic or pragmatic categories. The paper first introduces available corpora, typical research questions and methods in Interactional Linguistics. Then, two examples from the author's own research on verbal pseudo-coordination in German (*Sie können sich nicht hinstellen und sagen: »Das geht nicht.«* – *→You cannot stand up and say: »That is not okay.«*) are given to show which interpretive challenges the coding and quantification prosodic and pragmatic categories offer. Finally, the usefulness of the concept *hermeneutics* for the interpretive processes that the article describes will be discussed.

Keywords Interactional Linguistics · Categorization · Coding · Corpus · Pseudo-Coordination · Quantification

1 Einleitung

Die Interaktionale Linguistik (vgl. Couper-Kuhlen/Selting 2018; Imo/Lanwer 2019) ist ein Forschungsansatz, der erst in den letzten beiden Jahrzehnten zunehmend ein im eigentlichen Sinne korpusbasierter Ansatz geworden ist. Die Ursprünge des Ansatzes liegen in der Konversationsanalyse (CA) (vgl. Birkner et al. 2020; Sidnell/Stivers 2013), die sich der Ethnomethodologie zuordnet und traditionell kleine Belegkollektionen eines Phänomens handlungsanalytisch auswertet. Aus eher auf die sprachliche Form fokussierten konversationsanalytischen Arbeiten hat sich vor rund 25 Jahren dann der Ansatz der Interaktionalen Linguistik (IL) entwickelt. IL-Studien untersuchen meist syntaktische oder prosodisch-phonetische Formate in Bezug auf ihre konversationellen Funktionen. Die Datengrundlage der Untersuchungen waren zunächst meist Gesprächsaufnahmen, die einzelne Forschende oder Forschungsgruppen zu spezifischen Untersuchungszwecken selbst gemacht haben. Diese kleinen Korpora waren in der Regel nicht nach bestimmten Standards aufbereitet und nicht breit wiederverwendbar. In den vergangenen 15 Jahren wurden dann vermehrt Datenbanken und Korpora aufgebaut, bei denen von vornherein die (Wieder-)Nutzung zu unterschiedlichen Zwecken vorgesehen war. Das größte mündliche Referenzkorpus des Deutschen ist derzeit das Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK, vgl. Schmidt 2017b). Es wird von Forschenden für viele unterschiedliche Fragestellungen genutzt; in den meisten Fällen werden dabei die dem Korpus entnommenen Daten gemäß der jeweiligen Forschungsfrage hinsichtlich re-

relevanter Merkmale kodiert. Der Prozess des Kodierens ist zentral sowohl für die korpuslinguistische als auch für die interaktionslinguistische Auswertung. Kodieren bedeutet, Daten systematisch und nach eindeutigen, dokumentierten Kriterien zu kategorisieren. Mit der Kodierung wird die Grundlage gelegt für nachprüfbar qualitative und auch quantitative Auswertungen von Kategorien und Zusammenhängen zwischen diesen. Im Folgenden soll aufgezeigt werden, welche Notwendigkeiten, Möglichkeiten und Schwierigkeiten das Kodieren von Korpusdaten für die IL mit sich bringt. Dazu wird zunächst auf die vorhandene Strukturierung und Aufbereitung der Daten im Korpus anhand des Beispiels FOLK eingegangen. Anschließend werden unterschiedliche Arten von IL-Fragestellungen thematisiert, die mit unterschiedlichen methodischen Erfordernissen einhergehen, und ein Überblick über den Metadiskurs zum Kodieren in CA/IL gegeben. Schließlich wird anhand eigener Studien exemplarisch gezeigt, welche Möglichkeiten und Schwierigkeiten das Kodieren besonders für die Mündlichkeit relevanter Kategorien mit sich bringt und wie die entsprechenden Auswertungen und Interpretationen aussehen können. In Bezug auf die hermeneutische Unterscheidung von Verstehen und Interpretation steht hier letztere im Fokus: Sowohl das Kodieren selbst als auch das Deuten der Auswertung kodierter Kategorien sind interpretative Verfahren. Der Begriff Hermeneutik selbst ist in der CA/IL unüblich. Da er eine Vielzahl von Bedeutungen hat, die nicht alle gut auf die hier beschriebene Methodologie anwendbar sind, würde es eigentlich einer spezifischen, gut begründeten Definition von Hermeneutik bedürfen, die hier vorweg nicht geleistet werden kann.

Es wird aber in den Schlussreflexionen ein Bezug zur Hermeneutik-Literatur hergestellt.¹

2 Korpora in der IL

2.1 Zur Aufbereitung von FOLK

Am Leibniz-Institut für Deutsche Sprache (IDS) in Mannheim werden Korpora archiviert, die zwar zu einem großen Teil nicht für konversationsanalytische Zwecke entstanden sind, sondern z. B. zur Dokumentation von Varietäten des Deutschen dienen, die aber durchaus auch für IL-Studien nutzbar sind. Diese Korpora werden

¹ Auch *Interpretieren* bzw. *Interpretation* wird mit unterschiedlichen Bedeutungen verwendet. Meine Verwendung von Interpretation aus Sicht der forschenden Person entspricht der von Müller (2024), der Interpretation mit Dilthey – in hermeneutischer Tradition – als »regelgeleitetes Verstehen« und individuellen Prozess fasst. Aus Sicht der Forschenden muss »regelgeleitet« m. E. bedeuten, dass es um explizite oder explizierbare, begründete Regeln geht, wobei diese nicht notwendigerweise (sub-)disziplinweit übereinstimmend angewandt werden müssen. Im Gegensatz dazu kann Interpretation aus Sicht der Sprecher:innen bzw. im alltagssprachlichen Sinne auch durch unbewusste, implizite Regeln geleitet sein (und damit auch, enger gefasst, als *Verstehen* bezeichnet werden). Diese kommen in der IL/CA zwar beim Nachvollzug der Interpretationen der Sprecher:innen ins Spiel, bei den im vorliegenden Artikel fokussierten Kodierkategorien geht es aber um die linguistischen Interpretationen der Forschenden. Bei diesen spielt natürlich auch immer (fachliches) Vorwissen eine Rolle, sowohl bei theoriegeleitetem als auch bei datengeleitetem Vorgehen (vgl. Bender 2024), was reflektiert werden sollte.

vom Archiv für Gesprochenes Deutsch (AGD; vgl. Reineke/Schmidt 2022) gepflegt.² Ein Teil der archivierten Korpora ist online über die Datenbank für Gesprochenes Deutsch (DGD; vgl. Schmidt 2017a) zugänglich. Teil der DGD ist auch das einleitend bereits erwähnte Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK; vgl. Schmidt 2017b).³ Anders als die anderen Korpora in der DGD wird dieses Korpus nicht nur archiviert, sondern es wurde gezielt als wachsendes Referenzkorpus konzipiert.

Das Forschungs- und Lehrkorpus für Gesprochenes Deutsch wird seit 2008 am Leibniz-Institut für Deutsche Sprache aufgebaut. Das Ziel des FOLK-Korpus ist die wissenschaftsöffentliche Bereitstellung einer großen, nach aktuellen Standards erschlossenen und breit diversifizierten Datenbasis zur Untersuchung gesprochener Sprache in natürlicher Interaktion. Zielgruppe des Korpus sind Forschende, Lehrende und Studierende aus Gesprächsforschung bzw. Konversationsanalyse und Interaktionaler Linguistik sowie aus Korpuslinguistik und angrenzenden Fachgebieten (Reineke/Deppermann/Schmidt 2023, S. 72).

FOLK enthält derzeit 347 Stunden an Audio- und Videoaufnahmen (Version 2.20, Juni 2023). Die zugehörigen Transkripte umfassen 3,3 Millionen laufende Wörter. Das Korpus ist auf drei Ebenen annotiert; dies umfasst eine orthographische Normalisierung, eine Lemmatisierung und ein für Gesprächstranskripte angepasstes *Part-of-Speech-Tagging* (zu letzterem siehe Westpfahl 2020). Damit ist eine präzise formbasierte Suche möglich, die sich von der Suche in Schriftkorpora vor allem dadurch unterscheidet, dass keine Satzgrenzen einbezogen werden können. Eine automatische syntaktische Segmentierung ist nicht möglich, da die Transkripte keine orthographischen Satzzeichen enthalten.⁴ Es gibt aber eine Positionssuche, mit der sich auf Redebeiträge bezogene Einschränkungen hinsichtlich der Stellung eines Phänomens in die Abfrage einbeziehen lassen.

FOLK enthält eine Vielzahl von informellen und formellen Gesprächstypen (von Tischgesprächen und privaten Telefonaten über berufliche Meetings und soziolinguistische Interviews bis hin zu universitären Prüfungen und Podiumsdiskussionen). Die zu den Gesprächen gehörigen Metadaten (z. B. Jahr und Ort der Aufnahme, Alter und Herkunftsregion der Sprecher:innen) können in die Suche einbezogen werden (vgl. Deppermann/Reineke 2023). Zum Nutzen von FOLK für IL-Fragestellungen und zum allgemeinen Vorgehen dabei (Suche, Kollektionsbildung, (Re-)Kodierung, Analyse/Auswertung, Ergebnisdarstellung) sind einige Handreichungen und Artikel (vgl. z. B. Kaiser 2016; Kaiser/Schmidt 2016; Reineke/Deppermann/Schmidt 2023) erschienen, so dass hier auf eine diesbezügliche detaillierte Darstellung verzichtet wird.

² Zur Geschichte des Archivs siehe <https://agd.ids-mannheim.de/geschichte.shtml> sowie Stift/Schmidt (2014).

³ Siehe auch <https://agd.ids-mannheim.de/folk.shtml>.

⁴ Eine manuelle syntaktische Annotation von FOLK-Daten wurde im Projekt SegCor (<https://segcor.cnrs.fr/>) erprobt. Diese ist an das topologische Satzmodell angelehnt (vgl. Westpfahl et al. 2018, 2019). Die manuelle Annotation erwies sich als zu zeitaufwändig, um über das Pilotkorpus hinaus für FOLK insgesamt bereitgestellt zu werden.

2.2 Andere Korpora

Andere große Korpora und Datenbanken, die deutschsprachige Interaktionen enthalten, sind größtenteils nicht im selben Maße wie FOLK wissenschaftsöffentlich zugänglich (zu einem Überblick siehe Schmidt 2018, 2022). So ist beispielsweise die umfangreiche linguistische Audiodatenbank IAUda nur Germanistik-Studierenden und Lehrenden der Universitäten Hamburg und Münster zugänglich.⁵ Ausnahmen zur eingeschränkten Zugänglichkeit sind zwei komparative Korpora, die auch deutsche Daten enthalten: GeWiss, ein Korpus zur gesprochenen Wissenschaftssprache (Deutsch, Englisch, Polnisch, Italienisch), ist ähnlich aufbereitet wie FOLK und die deutschen Daten sind auch bereits Teil der DGD (vgl. Fandrych/Meißner/Wallner 2017). Auch Daten des noch im Aufbau befindlichen Korpus PECII (*Parallel European Corpus on Informal Interaction*; Deutsch, Englisch, Italienisch, Polnisch, Französisch, Finnisch) sollen Teil der DGD werden und werden von vornherein entsprechend aufbereitet (vgl. Kornfeld/Küttner/Zinken 2023). Diese Korpora sind nicht nur aufgrund ihrer deutschen Daten eine wachsende Ergänzung zu FOLK, sondern sie ermöglichen durch ihre komparative Ausrichtung, bei der über unterschiedliche Sprachen hinweg Gesprächs- und damit Handlungstypen möglichst konstant gehalten werden, Untersuchungen im Rahmen einer sprachvergleichenden CA/IL bzw. *pragmatic typology*, die in jüngster Zeit international stark betrieben wird (vgl. z. B. Dingemanse/Enfield 2015; Floyd/Rossi/Enfield 2020b).

3 Zu verschiedenen IL-Fragestellungen

Es lassen sich grob zwei für die CA/IL typische Arten von Fragestellungen unterscheiden: funktionsbasierte und formbasierte Untersuchungen, also z. B. »Wie werden im Deutschen Antworten auf Entscheidungsfragen gegeben?« (funktionsbasiert) vs. »Welche Funktion(en) hat die Responsivpartikel *nee* im Deutschen?« (formbasiert). Bei funktionsbasierten Untersuchungen ist in der Regel keine maschinelle Suche möglich, da sprachliche Handlungen nicht in Korpora annotiert sind und sich auch nicht eindeutig an bestimmten Formen festmachen lassen. Deshalb müssen Aufnahmen manuell nach dem interessierenden Phänomen durchgesehen und Fälle gesammelt werden. Die so gebildete Kollektion wird dann analysiert, indem Einzelfälle verglichen und nach bestimmten rekurrenten Merkmalen kategorisiert werden. Die relevanten Merkmale können die grammatische, lexikalische und prosodische Realisierung ebenso betreffen wie die Sequenzposition oder sprachexterne Aspekte wie den epistemischen Status. Beispiele für Funktionen bzw. Handlungen, die bereits als Basis für Untersuchungen gedient haben, sind Vorwurfsaktivitäten (Günthner 2000), *recruitments* (Floyd/Rossi/Enfield 2020b; Kendrick/Drew 2016) und Reparaturen (vgl. z. B. Dingemanse/Enfield 2015; Egbert 2009). Die untersuchten Handlungskategorien können dabei aus Vorannahmen oder anderen Untersuchungen stammen (z. B. Reparaturen) und dabei vorausgesetzt sein, aber

⁵ In der Regel muss ein eigener Beitrag zur Datenbank geleistet werden, um einen Zugang für die Arbeit an einer Qualifikationsarbeit zu bekommen (siehe <https://lauda.spracheinteraktion.de>).



einer Überprüfung unterzogen werden, oder aber sie können aus neuen Beobachtungen entstanden sein und anhand der Datenanalysen konsolidiert werden (z. B. *recruitments*).⁶

Bei formbasierten Untersuchungsfragen kann dagegen (je nach Abstraktheitsgrad des Phänomens und Aufbereitung des Korpus) eine maschinelle Suche der Ausgangspunkt sein.⁷ In der Regel muss zur Erstellung der Kollektion das Suchergebnis manuell durchgesehen werden, um zu entscheiden, welche Belege tatsächlich das interessierende Phänomen enthalten und welche nicht. Zum Beispiel müssten nach einer Suche nach der Wortfolge *sagen wir*, die auf das Auffinden des entsprechenden pragmatischen Markers abzielt, Belege ausgeschlossen werden, nach denen zwischen den beiden gesuchten Wörtern eine syntaktische Grenze liegt (*Er kann dann sagen: »Wir machen das nächstes Jahr.«*). In manchen Fällen muss anschließend auch noch eine Auswahl nach semantischen oder pragmatischen Kriterien erfolgen, z. B. wenn es formgleiche Phänomene mit deutlich anderer Funktion gibt. Formbasierte Untersuchungen gibt es z. B. zu Imperativen (Zinken/Deppermann 2017), Pseudoclefts (Günthner/Hopper 2010), rechtsversetzten Komplementsätzen (Proske/Deppermann 2020), *was heißt X* und *du meinst X* (Helmer/Zinken 2019) und anderen Konstruktionen mit bestimmten Verben (vgl. Deppermann/Proske/Zeschel 2017). Bei diesen Untersuchungen werden grammatische Kategorien vorausgesetzt und teilweise aber auch aufgrund der Variationsbreite der Formen in der Mündlichkeit neu abgesteckt. Wie bei funktionsbasierten Studien werden auch hier Zusammenhänge von Form und Funktion eruiert. In funktionsbasierten Untersuchungen wird die formale Bandbreite einer Handlung beschrieben, während in formbasierten Untersuchungen die funktionale Bandbreite einer Form beschrieben wird. Manchmal werden auch unterschiedliche Formen mit ähnlicher Funktion verglichen und hinsichtlich ihrer formalen und funktionalen Spezialisierung beschrieben. Ein Beispiel ist die Studie von Helmer/Zinken (2019), die qualitativ und quantitativ zeigt, dass *das heißt* eher für die Einleitung von *»(Re-)Formulierungen«* von Partneräußerungen verwendet wird, und zwar meist in Satzform, während *du meinst*, meist gefolgt von einer nicht satzwertigen Phrase, eher Reparaturen initiiert. Das Korpus FOLK ist zwar für verschiedene IL-Fragestellungen nutzbar, die meisten Untersuchungen, die es nutzen, sind aber formbasiert (vgl. Reineke/Deppermann/Schmidt 2023, S. 79).

⁶ Weitere Beispiele zu funktionsbasierten Untersuchungen (anhand von FOLK-Daten) finden sich in Reineke/Deppermann/Schmidt (2023, S. 79).

⁷ Eine maschinelle Suche ist möglich, wenn das interessierende Phänomen mindestens ein lexikalisch fixes Element enthält, das sich in seiner Variationsbreite durch die Annotationsebenen des Korpus finden lässt (z. B. *was heißt X*). Wenn die lexikalisch fixen Elemente allerdings zu häufig sind, ergibt eine maschinelle Suche meist vorwiegend Treffer, die nicht der gesuchten Konstruktion entsprechen, so dass man wie bei einer funktionsbasierten Untersuchung auf die Bildung eines Teilkorpus, das man manuell durchsieht, angewiesen ist. Ein Beispiel sind rechtsversetzte Komplementsätze, die sich nur an den beiden extrem häufigen Wörtern *das* und *dass* festmachen lassen (*das ist toll, dass du kommst*, vgl. Proske/Deppermann 2020).

4 Kodieren und Auswerten in der Interaktionalen Linguistik

Die CA und die IL gehen – zumindest traditionell – zunächst ethnomethodologisch vor. Das heißt, dass sie versuchen, aus den Daten heraus und aus der Perspektive der Interaktionsteilnehmenden relevante Kategorien zu rekonstruieren:

Erstens will sich die Konversationsanalyse auf das unmittelbar Beobachtbare beschränken und all ihre Aussagen *in den Daten* verankern. Zweitens stellt sie einen *streng rekonstruktiven Anspruch*: Es interessiert nicht, wie ein Analytiker Gesprächsaktivitäten aufgrund seiner Intuitionen oder theoretischen Ausrichtung versteht. Es geht vielmehr darum zu rekonstruieren, wie die Gesprächsteilnehmer selbst einander verstehen und an welchen Regeln oder Prinzipien sie sich dabei orientieren (Deppermann 2000, S. 98 f., Hervorhebung im Original).

Für diese Rekonstruktion orientieren sich Analytiker:innen an Handlungen, mit denen die Interaktionsteilnehmenden einander – und damit über das Transkript auch Außenstehenden – anzeigen, dass und wie sie einander verstanden haben. Ein einfaches Beispiel für dieses *display*-Prinzip: Reagiert eine Person auf den Turn einer anderen Person mit dem Responsiv *ja*, so zeigt sie an, dass der Vorgängerturn als Entscheidungsfrage verstanden wurde. Problematisch ist zum einen, dass es Kategorien gibt, die entweder (fast) nie von den Interagierenden selbst aufgezeigt werden (z. B. grammatische Kategorien) oder aber in der Regel nicht angezeigt werden (z. B. Verstehen). So folgert auch Deppermann (2000, S. 99): »[Die *display*-Konzeption] verkennt, dass das Erfassen von Sinn und Ordnung im Gespräch keine Frage eines einfachen Ablesens oder Hörens ist. Das Verständnis von Sinn und Ordnung bedarf vielmehr der Interpretation.« Und so fließen in der Praxis der CA natürlich doch viel sprachliches und außersprachliches Wissen der Forschenden ein, was nicht immer, aber doch zunehmend, methodisch reflektiert wird (vgl. auch Deppermann 2012). Insbesondere in der deutschsprachigen IL gibt es mittlerweile nicht nur einige *mixed-methods*-Ansätze, die statistische und damit per se analytikergebundene Methoden einbeziehen (siehe unten), sondern auch Forschende, die sich auf gebrauchsbasierte und kognitive Grammatiktheorien beziehen und damit neben Teilnehmerkategorien immer auch reine Analytikerkategorien verwenden (vgl. z. B. Ehmer 2011, 2022; Lanwer 2018b, 2020; Zima 2013). Auch meine eigenen Arbeiten greifen auf quantifizierende Methoden und grammatiktheoretische Kategorien zurück (siehe 5).

Das Kodieren von Daten, also das formale Auszeichnen von Fällen für (Ausprägungen von) Kategorien, für die man eindeutige Kriterien bzw. Indikatoren festgelegt hat, gehört in der korpusbasierten Linguistik ebenso zum alltäglichen Geschäft wie in anderen Forschungsgebieten.⁸ Die Konversationsanalyse jedoch hegt traditionell

⁸ Kodieren entspricht manuellem Annotieren, wie es z. B. von Müller (2024) beschrieben wird. Der Grund für die Tendenz zur Verwendung der Bezeichnung *Kodieren* in der IL und CA ist vermutlich, dass die Bezeichnung *Annotieren* mit Automatisierung, also entweder maschinell oder maschineller Nutzung manuell erstellter Annotationsschemata, assoziiert wird – Verfahren, die in diesen Forschungsrichtungen nicht verbreitet sind. Kodieren dient der Kategorienbildung und kann Voraussetzung für qualitative und quantitative Auswertungen sein. Vgl. auch Bender (2024) zu Kodierverfahren als induktive, iterative Prozesse der Kategorienbildung in den Sozialwissenschaften.



eine Skepsis gegenüber dem Kodieren und Quantifizieren (vgl. z. B. Schegloff 1993). Was das Kodieren angeht, mag dies zunächst verwundern, da natürlich auch bei CA-Analysen von Einzelfällen und Kollektionen Kategorien zum Einsatz kommen, z. B. wenn ein Turn(-Bestandteil) als Frage, Antwort, Reformulierung, Vorschlag, Ablehnung usw. eingeordnet wird. Diese »informelle« Art des Kodierens (vgl. Steensig/Heinemann 2015; Stivers 2015) geht jedoch immer mit einer detaillierten qualitativen Beschreibung weiterer Merkmale und Kategorien, die für den Einzelfall – insbesondere in Zweifelsfällen, die sich schwer einer Kategorie zuordnen lassen – relevant sind, einher. Für »formelles« Kodieren gilt: »virtually any code drastically reduces and bleaches the cases compared with what a moment-by-moment analysis can offer.« (Stivers 2015, S. 14)

Die informellen CA-Kodierungen werden in der Regel nicht in (statistische) Quantifizierungen überführt, wohl aber in Generalisierungen (»distributional evidence«, Stivers 2015, S. 5f.), wenn z. B. festgestellt wird, dass ein bestimmtes sprachliches Format »meistens« oder »häufig« eine bestimmte Handlung (z. B. Ablehnung eines Vorschlags) ausführt. Wie Stivers (2015, S. 6ff.) ausführt, wäre es nun aber nicht ausreichend für eine CA-Argumentation, solche informellen Quantifizierungen einfach in formelle statistische Auswertungen zu überführen, weil ein statistisches Ergebnis allein noch keine hinreichende Evidenz für die Orientierung der Sprecher:innen an so erkennbaren Präferenzen ist:

(...) distributional evidence is important to an argument that there is a social norm or that there is a preference for a given practice, but it is insufficient evidence without the additional support of speaker and recipient orientations to the practice and an understanding of cases that do not follow the generally observed pattern (Stivers 2015, S. 8).

Eine Sprecher:innenorientierung kann nur in qualitativen Analysen gezeigt werden. Sie ließe sich natürlich ebenfalls kodieren und quantifizieren, was jedoch wenig sinnvoll ist, weil sie üblicherweise in den allermeisten Belegen gerade nicht angezeigt wird, sondern eher selten an Reparaturen oder Thematisierungen erkennbar wird. Manche Fragestellungen sind aber insgesamt nur über eine stringente Quantifizierung bearbeitbar, vor allem wenn sie sehr viele, auch sprachexterne Variablen betreffen und wenn die Datenmenge insgesamt sehr groß ist (vgl. a. Mundwiler et al. 2019; Stivers 2015). Deshalb plädiert Stivers (2015) trotz dreier Problemfelder, die sie beschreibt (scharfe Kategoriegrenzen, Abgeschlossenheit des Kodierschemas, potenziell unangemessene Nutzung von *Mixed-methods*-Ansätzen), für eine Öffnung der CA hin zu Kodierungen und Quantifizierungen.

Die Stärke von CA/IL-Arbeiten ist es, dass Kodierkategorien auf qualitativen Analysen aufbauen; es wird versucht, die beobachteten Kategorien geeignet abzubilden. Somit sind sie empirisch gerade im funktionalen Bereich oft besser begründet und differenzierter als manche korpuslinguistische Untersuchungen, die formale Kategorien nicht immer funktional binnendifferenzieren (können). Steensig/Heinemann (2015) betonen, dass Kodierungen auch besonders als Anstoß für neue qualitative Untersuchungen gesehen werden sollten, weil man oft beim Kodieren auf zuvor nicht im Bewusstsein stehende Kategorisierungsprobleme stößt. Häufig heben Reflexionen zum Thema Kodieren in CA/IL auch hervor, dass es ein iterativer Prozess

ist und das Kodierschema sich im Laufe der Analysen fortwährend verändert, weil man merkt, dass die bisherigen Kategorien oder Werte noch nicht für alle Belege deskriptiv passen und so Teile der Kollektion rekodiert werden müssen (vgl. z. B. Küttner/Kornfeld/Zinken 2023; Lanwer 2020, S. 245; Mundwiler et al. 2019, S. 373). In einigen Arbeiten wird auch der Nutzen von Visualisierungen, die auf der Grundlage formal kodierter Daten möglich sind, stark gemacht (vgl. z. B. Mundwiler et al. 2019, S. 367 ff.). In Lanwers (2018a, b, 2020) Ansatz der rekonstruktiven, statistisch fundierten Kollektionsanalyse geht es darum, die übliche IL-Kategorienbildung transparent zu machen, indem die Kodierung zahlreicher Merkmale bzw. Merkmalsausprägungen einer Netzwerkanalyse unterzogen wird und so das Maß der Ähnlichkeiten unter den Einzelfällen visualisiert wird (vgl. Lanwer 2018b, S. 243). Aus linguistischer Sicht werden durch diesen Ansatz außerdem weniger schon vorab feststehende Kategorisierungen in die Analyse hineingetragen, weil sich (Sub-)Kategorien aus den aus der Netzwerkanalyse hervorgehenden Ähnlichkeitsrelationen erst ergeben und zudem nicht als scharf voneinander abgegrenzt, sondern als prototypisch organisiert konzeptualisiert werden.

In jedem Fall kann das Kodieren und Quantifizieren mittlerweile als in Teilen der CA/IL etabliert gelten, wovon auch die wachsende Zahl an veröffentlichten Kodierschemata/-manualen zeugt, der vor allem im Bereich der sprachvergleichenden CA/IL vorliegt (vgl. Dingemanse/Kendrick/Enfield 2016; Floyd/Rossi/Enfield 2020a; Küttner/Kornfeld/Zinken 2023; Stivers/Enfield 2010).

Viele allgemeine Schwierigkeiten, die im Laufe von Kodierprozessen bearbeitet werden müssen, teilen IL-Arbeiten mit korpuslinguistischen Untersuchungen. Dies sind z. B. die Entscheidung für eine Granularitätsebene, der Umgang mit Zweifelsfällen, Doppel-Kodierung im Hinblick auf Ausprägungen einer Kategorie (wird eher gemieden), Doppel-Ratings (bevorzugt, nicht immer gemacht).⁹ Besonderheiten liegen im Hinblick auf die Gewinnung der zu kodierenden Datenbasis vor: Da mündliche Korpora im Vergleich zu Schriftkorpora klein sind und viele Phänomene darin nicht häufig genug vorkommen, kann nicht immer eine zufällige Stichprobenziehung erfolgen, sondern oft müssen gezwungenermaßen exhaustiv alle Belege eines Phänomens genutzt und kodiert werden, ohne dass immer für alle Varianten des Phänomens ausreichend Fälle vorliegen (vgl. auch 5.2.2).

5 Ein Beispiel

5.1 Das Phänomen: Pseudokoordination

In diesem Abschnitt wird ein Phänomen aus meiner Forschung fokussiert, mit dessen Untersuchung ich schon vor zehn Jahren begonnen habe, das aber aufgrund seiner Komplexität und Struktur und auch aufgrund der Datenlage (was geringe Menge und große Variationsbreite betrifft) immer wieder neue Teiluntersuchungen

⁹ Bisher wird in nur wenigen Publikationen aus dem Bereich der IL über das Kodieren von Belegen durch zwei Personen berichtet und noch seltener werden Maßzahlen fürs Inter-Annotator-Agreement angegeben (vgl. z. B. Helmer/Zinken 2019; Zinken/Küttner 2022).

und -kodierungen erfordert. Damit eignet es sich besonders, einige diesbezügliche Aspekte zu illustrieren.

Die grammatische Konstruktion, um die es geht, wird oft als Pseudokoordination (PK) bezeichnet (zu Begrifflichkeit und sprachübergreifender Literatur siehe Proske 2017, 2019, 2023). Es handelt sich um eine Koordination von zwei Verben (oder, je nach Definition, von zwei Verbalphrasen), die semantisch gesehen nicht gleichrangig sind (deshalb *pseudo-*). Sie stehen in einer im weitesten Sinne kausalen oder finalen Relation zueinander. So wird z. B. dann, wenn die Bewegungsverben *kommen* und *gehen* im ersten Konjunkt stehen (V1), durch das Verb im zweiten Konjunkt (V2) der Zweck des Ortswechsels verbalisiert:

- (1) gustaf **kommt** MORgen bei uns **und**: (0.2) **macht** die RITzen sauber.
(FOLK_E_00201_T_01 / c282)

Beispiel 1¹⁰ ist annähernd gleichbedeutend mit *Gustav kommt am Montag, um die Ritzen sauberzumachen*. Das Verb im zweiten Konjunkt ist also semantisch untergeordnet, auch wenn es die eigentlich relevante Handlung beinhaltet. Auch ohne das erste Konjunkt wäre klar, was passiert: *Gustav macht am Montag die Ritzen sauber*. Dass er dafür kommen muss, ist impliziert, wenn bekannt ist, dass der Subjektreferent nicht am zu putzenden Ort wohnt. Im ersten Konjunkt von Pseudokoordinationen dagegen treten einige Verben und Verbklassen besonders häufig auf und tendieren zum Verlust ihrer literalen Bedeutung und zur Entwicklung aspektueller und/oder subjektiver Bedeutungskomponenten. Dies verdeutlicht das folgende Beispiel aus FOLK; Ministerin Theresia Bauer äußert sich in einer Podiumsdiskussion zur Situation der Musikhochschulen in Baden-Württemberg:

- (2) wir **HABen** eine ungleichverteilung-
°h ich **kANN** aber beim **bESTen** willen jetzt nicht **HERgehn**-
°h **und Anfangen** die musikhochschulen **gLEIchmäßig** über baden
WÜRTtemberg zu verteilm-
(FOLK_E_00126_T_02, c606)

Das Verb *hergehen* hat hier keine Bewegungsbedeutung, sondern drückt eine Haltung des Subjektreferenten (›Intentionalität‹; ›Entschlossenheit‹) aus. Diese Haltung wird dem Subjektreferenten von der Sprecherin zugeschrieben;¹¹ im vorliegenden Beispiel wird diese Haltung und die Handlung, die mit dieser Haltung durchgeführt wird (*anfangen, die Musikhochschulen gleichmäßig über Baden-Württemberg*

¹⁰ Alle Beispiele sind nach dem Gesprächsanalytischen Transkriptionssystem GAT 2 (vgl. Selting et al. 2009) transkribiert. Jede Zeile stellt eine Intonationsphrase dar, die mit einer finalen Tonhöhenbewegung endet (Punkt: stark fallend; Semikolon: leicht fallend; Bindestrich: gleichbleibend; Komma: leicht steigend; Fragezeichen: stark steigend). Silben mit Fokusakzent sind in Großbuchstaben wiedergegeben, in Silben mit Nebenakzent ist nur der Vokal groß gesetzt. In den Beispielen wurden die Sprechersiglen weggelassen, weil die gewählten Beispiele nur Redezüge eines einzelnen Sprechers/einer einzelnen Sprecherin und somit keine Sprecherwechsel enthalten.

¹¹ Im Beispiel ist die Sprecherin mit dem Subjektreferenten identisch. Dies kommt zwar regelmäßig vor bei PK, in der Mehrheit der Fälle ist der Subjektreferent aber eine dritte Person, wie in Beispiel (3).

zu verteilen), abgelehnt (*ich kann beim besten Willen nicht*). In anderen Fällen wird die entschlossene Haltung, wie sie z.B. der Fahrlehrer in (3) durch *hingehen* ausgedrückt und auf die Handlung im zweiten Konjunkt (*Leistung zeigen*) bezieht, positiv bewertet.

- (3) DU willst den FÜHrerschein.
 ich HAB den.
 dU krist den NUR,
 (0.39) wenn du **HIN**gehst,
 °h **und** (.) LEISTung **zeigst**,
 indem du dich HOCHKonzentriert,
 (0.35) fließend in den verKEHR hineinbringst.

(FOLK_E_00146_T_02, c203)

Hier liegen also subjektive, insbesondere evaluative Bedeutungen im Sinne Traugotts (1988) vor. Traugott geht von einer Entwicklung solcher Bedeutungen aus nicht perspektivierten, ›objektiven‹ Bedeutung aus: »Meanings situated in the external described situation > meanings situated in the internal (evaluative/perceptual/cognitive) situation« (Traugott 1988, S. 409). Es fragt sich nun, woran man das Vorhandensein solcher Bedeutungen kodierbar festmachen kann. Die wertende Bedeutung von *her-/hingehen* ist nur bei Verwendung in der PK-Konstruktion vorhanden. Man könnte annehmen, dass eine subjektive Bedeutung in V1 lexikalisiert ist. Das ist aber nicht bei allen V1 so einfach, weil nicht in allen Belegen eindeutig eine wertende, nicht literale Bedeutung vorliegt. Eindeutig kodierbare Wertungen sind nur am verbalen Kontext zu erkennen, z.B. an der (negierten) Modalverbverwendung in (2) oder an wertenden Folgeformulierungen wie in (3) (*hochkonzentriert, fließend einbringen*). Zusätzlich zur Kodierung solch expliziter Indikatoren kann das Verblässen der V1-Bedeutung kodiert werden (und damit implizit oft, aber nicht immer, auch subjektive ›Aufladung‹, vgl. 5.2.2).

Die Datengrundlage aller meiner PK-Untersuchungen stammt aus FOLK-Suchen nach den Lemmata *kommen, gehen, sitzen* und *stehen* (und einigen anderen) mit *und* im rechten Kontext (im Abstand von bis zu 15 Tokens). Die erste Suche erfolgte in FOLK-Version 2.6 für Proske (2017). Bei jeder neu erschienenen Version wurde die Suchabfrage mit der Metadaten-Angabe ›in DGD seit Version X‹ wiederholt, um für Folgeuntersuchungen die aktuellste Datengrundlage vorauszusetzen. Zu den Einzelheiten der Auswahl der einschlägigen Belege aus der KWIC-Liste siehe Proske (2017, 2019, 2023).

5.2 Forschungsfragen, Kodierung, Auswertung und Interpretation

Es gibt in der Literatur zahlreiche Hypothesen und Forschungsergebnisse zur PK in anderen Sprachen, die in Bezug auf die Konstruktion im Deutschen noch zu testen waren oder sind. Im Folgenden sollen zwei spezifisch interaktionsbezogene Behauptungen aus der Literatur herausgegriffen werden, um allgemeine Verfahren und Schwierigkeiten bei der Kodierung und Auswertung von nicht-binären und nicht formal eindeutigen Kategorien zu diskutieren. Diese Kategorien sind zum einen die



Prosodie bzw. die Gliederung der pseudokoordinierten Sätze in Intonationsphrasen (IPn) und zum anderen die pragmatische Funktion, die eng mit dem semantischen Merkmal des Grads der Subjektivität zusammenhängt.

5.2.1 Prosodie als Indikator für Monoprädikativität oder Konstruktionsstatus

Die erste Forschungsfrage betrifft die für die Pseudokoordination konstitutive »Monoprädikativität« (Hopper 2008, S. 255), also die semantische Abhängigkeit der beiden Verben voneinander, und ihren möglichen Zusammenhang mit formalen Realisierungsmerkmalen. Viel diskutiert wird in der Literatur z. B., dass eine direkte Adjazenz von V1, Konjunktion und V2 für einen hohen Grad an konzeptueller Kohäsion spreche.¹² Bisher selten diskutiert wurde die prosodische Kohäsion als Indikator für semantische Einheit, wohl auch weil meist mit Schriftdaten gearbeitet wurde. Barth-Weingarten/Couper-Kuhlen (2011) haben als erste fürs Englische untersucht, ob pseudokoordinierte VPn eher in einer gemeinsamen Intonationsphrase realisiert werden als koordinierte VPn. Ihr Ergebnis war, dass Monoprädikativität tendenziell mit weniger prosodischen Grenzsignalen zwischen den Konjunkten einhergeht:

[...] a pilot study [...] showed significant correlations between monopredicativity and features of prosodic togetherness, among them lack of pitch reset with *and* and lack of prosodic breaks such as filled/unfilled pausing but presence of latching at the potential unit boundary. Those VP conjunctions which were interpretable in context as two events, in contrast, exhibited significantly more disintegrative prosodic/phonetic features [...] (Barth-Weingarten/Couper-Kuhlen 2011, S. 280)

Die Forschungsfrage nach dem Zusammenhang von prosodischer und semantischer Kohäsion betrifft (mindestens) zwei zu kodierende Kategorien:

- a) (Grad oder Vorhandensein der) Monoprädikativität
- b) (Grad oder Vorhandensein einer) prosodischen Grenze zwischen Konjunkt 1 und Konjunkt 2

Die eingeklammerte Formulierung »Grad oder Vorhandensein« macht deutlich, dass die Kategorien »Monoprädikativität« und »prosodische Grenze« entweder binär

¹² Im Deutschen gibt es oft trotz semantischer Kohäsion keine direkte Adjazenz, weil diese teilweise aufgrund der Wortstellungsregeln nicht möglich ist. Daher ist der Grad der syntaktischen Kohäsion möglicherweise nicht in jeder Sprache ein guter Indikator für konzeptuelle Kohäsion. Trotzdem kann sich das Auswerten der syntaktischen Kohäsion lohnen: Die Kodierung der Menge der zwischen die Verben und die Konjunktion tretenden Tokens hat im Falle von *kommen* als V1 gezeigt, dass Fälle mit *sagen* als V2 häufiger adjazente Stellung aufweisen als Fälle mit anderen V2 (vgl. Proske 2017). Dies lässt sich als Tendenz zur separaten Verfestigung der Verbindung *kommen und sagen* interpretieren. Allerdings wirft auch hier die Auswertung ähnliche Fragen wie die für die Prosodie diskutierten auf: Ist es für die Interpretation als syntaktisch kohäsiv entscheidend, dass sowohl zwischen V1 und die Konjunktion als auch zwischen die Konjunktion und V2 keine Wörter treten? Oder reicht eine direkte Adjazenz von V1 und Konjunktion, weil die grammatischen Regeln hier viel mehr Möglichkeiten zur Platzierung von Satzgliedern lassen als zwischen Konjunktion und V2?

(vorhanden vs. nicht vorhanden) oder als nicht-binär (anhand von mehr als zwei spezifischen Merkmalsausprägungen) kodiert werden können.

Barth-Weingarten/Couper-Kuhlen (2011, S. 280) verwenden in ihrer Untersuchung drei Ausprägungen für die Kategorie Monoprädikativität: »1 event«, »1-2 events«, »2 events«. Es wird nicht offengelegt, wie genau die Klassifikation erfolgt ist. Zwar finden sich in den Beispielbeschreibungen Hinweise, welche Indikatoren für Monoprädikativität herangezogen wurden, aber die Kodierung wird nicht genau beschrieben. Insbesondere wird nicht erläutert, aus welchen Gründen Belege in die recht große Gruppe (34 von 96) der Zweifelsfälle (»1-2 events«) eingeordnet wurden. Für meine Untersuchungen wurde Monoprädikativität folgendermaßen operationalisiert: Ist die Pseudokoordination in ein Haupt-Nebensatzgefüge mit vergleichbarer Bedeutung umformbar (siehe Beispiel 1 oben),¹³ so stellen die beiden Prädikate Teilhandlungen einer übergeordneten komplexen Handlung dar und werden als monoprädikativ=1 kodiert. Eine verblasste Verbsemantik ist demnach keine notwendige Voraussetzung für Monoprädikativität. Umgekehrt gilt aber: Liegt verblasste Verbsemantik vor, ist der Beleg monoprädikativ (zu weiteren Indikatoren, Einzelheiten und Schwierigkeiten vgl. Proske 2017, S. 189–193).

Wie ist nun die prosodische Gestaltung zu operationalisieren und kodieren? Barth-Weingarten/Couper-Kuhlen (2011: 280) verwenden für die Untersuchung der prosodischen Kohäsion zum einen eine binäre Unterscheidung in »no break« und »break« vor der Konjunktion. Zum anderen wird die Kategorie »break« intern ausdifferenziert anhand der Subkategorien »unfilled pause«, »breathing«, und »latching«. Diese Subkategorien sind einige der möglichen Indikatoren für prosodische Grenzen. Barth-Weingarten (2016) plädiert dafür, bei der Segmentierung von Gesprochenem nach Intonationsphrasen (beim Transkribieren und Analysieren) nicht nur klare Grenzen zuzulassen, sondern auch *Zäsuren* unterschiedlicher Stärke zu bestimmen. Liegen viele Zäsurierungsindikatoren zugleich vor, kommt das einer klaren Intonationsphrasengrenze nah. Liegt nur ein Indikator vor, kann die schwache Zäsur dennoch aus Analytiker- oder Sprecherperspektive relevant sein.

In meinen Kodierungen wird ebenfalls nicht binär unterschieden, sondern nach IP-Grenze (»1«), Zäsur (»Z«) und keiner Grenze/Zäsur (»0«) (vgl. Proske 2017, S. 208f.). Fälle ohne jede prosodische Grenze oder Zäsur zwischen den Konjunkten sind die Beispiele in (4). Sowohl *bis ich komm und den hol* als auch *und eines guten Tages kommt er und sagt* sind in einer IP mit nur einem Fokusakzent produziert.¹⁴ Eine Zäsur liegt in den Beispielen in (5) vor; hier ist jeweils auf dem Wort vor der Konjunktion *und* eine IP-finale Tonhöhenbewegung vorhanden (leicht steigend, symbolisiert durch das Komma).¹⁵ Die Beispiele in (6) zeigen klare IP-Grenzen zwischen den Konjunkten, die sich am Vorkommen mehrerer Indikatoren festmachen

¹³ Neben finalen/kausalen Relationen kann Pseudokoordination auch eine Gleichzeitigkeitsrelation herstellen: *ihr kommt grade REIN? (-) und seht eure VÄter* (FOLK_E_00330_T_02, c687) lässt sich mit einem temporalen Satzgefüge paraphrasieren: ›Als ihr reinkommt, seht ihr eure Väter.«

¹⁴ In 4a liegt der Fokusakzent auf V2 (*hol*), in 4b liegt er auf V1 (*kommt*). Das jeweils andere Verb trägt einen Nebenakzent (V1 *komm*; V2 *sacht*).

¹⁵ Es wurden bei der Kodierung als »Z« auch neben der Tonhöhenbewegung auch andere Zäsurierungsindikatoren einbezogen (z. B. Tempoveränderungen), auf die hier nicht näher eingegangen werden kann.

Tab. 1 Zusammenhang von Monoprädikativität und prosodischer Gestaltung von *kommen und V*

	klare Grenze	Zäsur	keine Grenze/Zäsur	Summe
monoprädikativ	22 (28 %)	21 (27 %)	35 (45 %)	78 (100 %)
nicht monoprädikativ	7 (41 %)	5 (29,5 %)	5 (29,5 %)	17 (100 %)
Summe	29 (31 %)	26 (27 %)	40 (42 %)	95 (100 %)

lassen; der finalen Tonhöhenbewegung folgt jeweils eine Pause, in 6a auch ein hörbares Einatmen (°h).

- (4) a. sollen se ihn schön lang AUFheben;
und von a nach BE räumen;
bis ich kOmm und den HOL.

(FOLK_E_00055_T_07, c234)

- b. **un eines guten tages KOMMT er un sacht-**
(.) da Oben de:r (.) müller
der (.) HOLT nachher das pony.

(FOLK_E_00342_T_01, c309)

- (5) a. isch hab gesagt **isch komm um VIER,**
un hol se AB.

(FOLK_E_00293_T_05, c641)

- b. **halbe stunde später KOMMT er,**
und sacht ach ich will DOCH nix.

(FOLK_E_00152_T_01, c64)

- (6) a. un wenn_s jetzt nich grad STOßzeiten sind,
kann sie ja dann in der zeit vor zwischen zehn un Elf KOMmen,
°h (.) und einen mit RÜbernehmen,

(FOLK_E_00082_T_01, c559)

- b. **da (0.51) kommt ja das eine KIND?**
(.) von von (.) CANDice?
(0.51) und sagt so zu isaBELla,
(0.31) oh;
so sieht tan äh tante isaBELla also aus.

(FOLK_E_00288_T_02, c69)

Für die Auswertung und Interpretation stellt sich nun die Frage, ob die Zäsur-Fälle den klaren Grenzen oder den Fällen ohne jeden Zäsurierungsindikator ähnlicher

sind. Für meine Untersuchung von *kommen und V* (Proske 2017) wurden 95 Belege ausgewertet. 40 davon wiesen eine klare IP-Grenze auf, 29 waren klar als nur eine IP realisiert, also ohne Grenze bzw. prosodisch integriert. 26 Fälle wurden als Zäsur kodiert. Schlägt man diese den integrierten Fällen zu, weil sie wie diese keine klare Grenze aufweisen, erhält man eine Mehrheit von 66 Fällen, die eher prosodisch integriert realisiert sind. Fasst man dagegen die Grenzen und Zäsuren zusammen, weil sie jeweils aus zwei IPn bestehen, erhält man eine Mehrheit von 69 Fällen, die eher desintegriert realisiert sind (Tab. 1).

Unabhängig davon, womit man die Zäsuren zusammenfasst, zeigt sich ein kleiner Unterschied: Monoprädikative Fälle sind häufiger prosodisch integriert und seltener desintegriert als nicht monoprädikative. Dies ist aber gar nicht besonders aussagekräftig, weil hier nicht monoprädikative Fälle mit klar nicht monoprädikativen Fällen, sondern nur mit semantischen Zweifelsfällen, die formal als pseudokoordiniert in Frage kommen, verglichen wird (anders als bei Barth-Weingarten/Couper-Kuhlen 2011¹⁶). Wichtiger ist es, sich zu überlegen, welchen Grad an prosodischer Integration man überhaupt bei der PK insgesamt im Deutsch erwartet, angesichts ihres geringen Grammatikalisierungsgrads und vor dem Hintergrund der Wortstellungsvarianz, und ob man nicht eher die prosodische Kohäsion von lexikalisch verfestigten Subkonstruktionen wie *kommen und sagen* mit dem Rest der PK-Belege mit dem entsprechenden V1 vergleicht. Insgesamt ist es wichtig, verschiedene V1, auch in Kombination mit frequenten V2, als separate Kollektionen oder Stichproben zu analysieren, anstatt von vornherein über PK mit allen möglichen unterschiedlichen Verben zu generalisieren. Wie die Belege 4b, 5b und 6b zeigen, folgt bei *sagen* als V2 in der Regel ein Komplementsatz, der das Gesagte wiedergibt. Bei anderen V2, die als Objekt eine NP nehmen, wie in 4a, 5a und 6a, steht das Objektkomplement eher vor dem zweiten Verb. Dies hat auch Einfluss auf die prosodische Gestaltung. Zum einen hat der Komplementsatz von *sagen* oft mehr prosodisches Gewicht als eine NP. Zum anderen beginnt die Redewiedergabe oft mit einer Vor-Vorfeldpartikel wie *ach*, *oh* oder *hallo*, die teilweise eine eigene prosodische Kontur hat und damit auch Anlass für Zweifelsfälle der prosodischen Segmentierung sein kann, die bei andern V2 und ihren Komplementen nicht vorkommen. In (7) finden sich zwei alternative Segmentierungen, je nachdem, ob man nach *kommt* oder nach *sagt*, die beide nur schwache Tonhöhenbewegungen aufweisen, eine IP-Grenze ansetzt oder nicht.

- (7) a. und DU stehst da in weITer flur allEIne,
 un die poliTESS kommt und sacht;
 haLLO,
 warum has dU zwei PARKplätze genommen.

¹⁶ Auch die prosodische Kodierung an sich ist zwar ähnlich, aber nicht hundertprozentig vergleichbar. Siehe auch Bergmann (2022) dazu, wie phonetische und funktionale Kategorien je nach Studie unterschiedlich operationalisiert werden und damit die Ergebnisse nicht unbedingt vergleichbar sind.

b. und DU stehst da in wEIter flur allEIne,
 un die poliTESS kommt-
 und sacht halLO,
 warum has dU zwei PARKplätze genommen.

(FOLK_E_00169_T_02, c671)

Bei Barth-Weingarten/Couper-Kuhlen (2011) erfolgt die Auswertung nicht nach verschiedenen V1 differenziert, ihre Interpretation der Befunde zielt auf prosodische Kohäsion als Voraussetzung für die Verfestigung der VP-*and*-VP-Konstruktion auf abstrakter Ebene. M.E. wäre eine *bottom-up* Argumentation, die von der Verfestigung von lexikalisch spezifischen Konstruktionen, mit bestimmten V1, ausgeht, plausibler. Ein Ansatz dazu findet sich bei Barth-Weingarten (2016, S. 255), wo in einer kleinen Pilotstudie 26 ältere und 20 neuere Belege von *go and* VP hinsichtlich ihrer prosodischen Gestaltung verglichen werden.

5.3 Indikatoren für subjektive Bedeutungsanteile

Die zweite Forschungsfrage betrifft die wiederholt in der Literatur erwähnte Tendenz zur Wertung, mit der die Pseudokoordination und verwandten Konstruktionen verbunden sind (vgl. Drew et al. 2021; Haddington/Janutunen/Sivonen 2011; Hopper 2002; Proske 2017; Stefanowitsch 2000). Die Arbeiten, die diese Tendenz beobachten, unterscheiden sich darin, woran die Wertung festgemacht wird und wie ihr Vorhandensein interpretiert wird. Die meisten Arbeiten gehen von einer Veränderung der Semantik von V1 aus, bei der bestimmte Bedeutungskomponenten, die an sich noch nicht wertend sind, z.B. Intentionalität, zu kontextuellen Inferenzen führen, so dass eine – oft, aber nicht notwendigerweise immer negative – Wertung mit der Zeit zum rekurrenten Merkmal wird (z.B. Entschlossenheit, Unverfrorenheit). Bei einigen Arbeiten wird die Semantik nicht unmittelbar als Voraussetzung für wertende Funktionen gesehen; es wird beschrieben, dass PK in negativ besetzten kommunikativen Kontexten vorkommt und damit assoziiert wird. Hier soll insbesondere Bezug auf letztere Behauptung genommen werden, weil sie eine spezifisch interaktionale Auswertung betrifft:

Our analysis of hendiadys in naturally occurring interactions shows that its use is systematically associated with talk about complainable matters, in environments characterized by conflict, dissonance, or friction that is ongoing in the interaction or that is being reported by one participant or another. We also find that the utterance in which hendiadys is used is typically in a subsequent and possibly terminal position in the sequence, summarizing or concluding it. (Drew et al. 2021, S. 323)

Das Zitat enthält zwei zu testende Hypothesen:

- a) Pseudokoordination kommt (vom Erstverb unabhängig) systematisch (=immer? meistens?) in konfliktbehafteten Kontexten vor und ist damit deutlich negativ konnotiert.

Tab. 2 Subjektive Wertung bei *sitzen/stehe*n und V

	neg. Kont.	neg. Subj.	pos. Kont.	pos. Subj.	keine Wertg.	Summe
<i>stehe</i> n	20 (30 %)	17 (25 %)	2 (3 %)	3 (5 %)	25 (37 %)	67 (100 %)
<i>sitze</i> n	24 (20 %)	18 (15 %)	8 (7 %)	9 (7 %)	61 (51 %)	120 (100 %)
Summe	44 (24 %)	35 (19 %)	10 (5 %)	12 (6 %)	86 (46 %)	187 (100 %)

b) Pseudokoordination tritt typischerweise (=immer? meistens?) in zweiter oder dritter Sequenzposition auf.

Hier soll nur erstere diskutiert werden. Drew et al. (2021) untersuchen je 20 Belege in vier Sprachen und demonstrieren in ihren Fallanalysen, dass alle Fälle in (unterschiedlichen) konfliktären Kontexten stehen. Eine Operationalisierung und Kodierung wurde nicht vorgenommen. Die Datenbasis, aufgrund derer generalisiert wird, ist vergleichsweise klein, zumal Fälle mit ganz unterschiedlichen V1 in die Kollektion eingehen.¹⁷ Vor diesem Hintergrund erscheint die Interpretation, die Drew et al. vorschlagen, recht weitreichend: Sie schreiben der Verwendung der Pseudokoordination in konfliktären Kontexten normativen Charakter bzw. Angemessenheit und Erwartbarkeit für das Berichten bedauerlicher Umstände zu (Drew et al. 2021, S. 340, 342). Sie machen dies daran fest, dass Korrekturen vorkommen, bei denen erst nur ein Verb verwendet wird und dann zur PK mit dem ursprünglich gewählten Verb als V2 gewechselt wird. Dass dies die Orientierung der Sprecher:innen am ›Mehrwert‹ von V1 zeigt, ist plausibel (vgl. auch Proske 2017, S. 193), aber was genau dieser Mehrwert ist, kann aus den Beobachtungen heraus unterschiedlich interpretiert werden. Dass er das Vorkommen in konfliktären Kontexten betrifft, ist ebenso eine Analytikerzuschreibung wie die Idee, dass eine bestimmte Semantik vermittelt wird, wie z.B. Intentionalität, die von Drew et al. (2021, S. 345, »imposition or deliberateness«) ebenfalls erwähnt wird. M.E. wäre also eine genauere Reflexion des Verhältnisses von Semantik und Pragmatik notwendig. Um der Bestimmung dieses Verhältnisses näher zu kommen, würde eine Kodierung und Quantifizierung beider Aspekte helfen, die zunächst zeigen könnte, welcher häufiger vertreten und damit unter Umständen auch für die Sprecher:innen primär ist. Ebenso wie Stivers argumentiert, dass quantitative Evidenz ohne Evidenz zur Sprecher:innenorientierung nicht hinreichend ist, kann man argumentieren, dass Evidenz zur Sprecher:innenorientierung in Kombination mit informeller distributioneller Evidenz nicht ausreichend ist, sondern eine formale, statistisch reliable Quantifizierung einbezogen werden muss.

Wie genau man sich nun einer Operationalisierung des Kontextes und der Bedeutung annähert, ist allerdings nicht trivial. Für meine Untersuchung von *sitze*n/*stehe*n und V (Proske 2023) habe ich für »Wertung« folgende Kodierkategorien verwendet: »negativer Kontext«, »positiver Kontext«, »negative Subjekthandlung«, »positive Subjekthandlung«, »keine«. Die folgenden Beispiele zeigen, dass der Subjektreferent sprecherseitig negativ (8) oder positiv (9) bewertet werden kann oder aber nur

¹⁷ Zudem wurden nur Belegen berücksichtigt, bei denen die beiden Konjunkte eine gemeinsame Intonationsphrase bilden (vgl. Drew et al. 2021, S. 324).

Tab. 3 Zusammenhang von literaler vs. verblasster V1-Bedeutung bei *sitzen/stehten* und V

	neg. Kont.	neg. Subj.	pos. Kont.	pos. Subj.	keine Wertg.	Summe
literale B.	19 (17%)	19 (17%)	7 (7%)	6 (6%)	58 (53%)	109 (100%)
Zweifelsfall	17 (31%)	7 (13%)	2 (4%)	6 (11%)	22 (41%)	54 (100%)
nicht-lit. B.	8 (33%)	9 (38%)	1 (4%)	0 (0%)	6 (25%)	24 (100%)
Summe	44 (24%)	35 (19%)	10 (5%)	12 (6%)	86 (46%)	187 (100%)

die Situation bzw. der Sachverhalt negativ (10) oder positiv (11) zu bewerten ist, der Subjektreferent aber nicht. Oft ist auch gar keine Wertung auszumachen (12).

(8) jemand der_s halt NICH kann,
der sitzt mit der WAFFE-
und macht einf macht einfach nur DAUerfeuer-
(FOLK_E_00182_T_02, c362)

(9) dis war do_mal die coole aktION wo ich-
n abend lang DAsaß un LIEDer reingeladen hab.
(FOLK_00165_T_02, c220)

(10) dann KOMMT er nich;
so.
und dann SITzen wir da und WARTen.
und dis is ÄTzend.
(FOLK_00285_T_02, c219)

(11) und diese kleinen °h öh öh sch STRAUßenwirtschafTen.
wo ma DRAUßen sitzt und äh-
°hh ne kleinigkeit ISST und n WEIN trinkt-
(.) dis HÄTT_ich gerne.
(FOLK_E_00421_T_01, c490)

(12) ach so wir ham da geSEssen und ZEITung gelesen.
(FOLK_E_00202_T_02, c39)

Die Beispiele geben hier nur kurze Ausschnitte wieder, für die Kodierung wurde aber ein größerer Kontext einbezogen. Die statistische Auswertung zeigt, dass negative Wertungen (des Kontexts und des Subjektreferenten zusammengenommen) zwar in einem großen Teil der Belege (55% für *stehten* und 35% für *sitzen*) vorhanden sind, während positive Wertungen eher selten sind (8% für *stehten* und 17% für

sitzen), dass aber ein großer Teil auch gar keine Wertung enthält (37 % für *stehen* und 51 % für *sitzen*) (Tab. 2).

Daraus lässt sich zunächst folgern, dass fürs Deutsche nicht davon ausgegangen werden kann, dass PK (nur) in negativ konnotierten Kontexten erwartbar ist. Überprüft man nun den Zusammenhang von negativer Wertung und verblasster Bedeutung,¹⁸ ergibt sich folgendes Bild: Bei literaler Bedeutung gibt es anteilig deutlich weniger Fälle von negativer Bewertung (34 %) als bei nicht-literaler Bedeutung (71 %) (Tab. 3).

Dies zeigt, dass verblasste Bedeutung und negative Wertung häufig miteinander einhergehen. Allerdings kommt eine klar verblasste Bedeutung überhaupt nur in 24 von 187 PK-Fällen (13 %) mit *sitzen* und *stehen* als V1 vor, was statistische Aussagen über diese Teilmenge wenig robust macht. Zudem bleibt es Interpretationssache, in welche Richtung man den Zusammenhang von Semantik und Pragmatik deutet: Da deutlich mehr von den Fällen mit negativer Wertung (erste und zweite Spalte) eine literale Bedeutung haben (38 von 79 = 48 %) als eine verblasste Bedeutung (22 %), könnte man argumentieren, dass die pragmatische Funktion primär ist. Andererseits kann eine nicht wertende PK (fünfte Spalte) deutlich schlechter durch nicht-liternale V1 ausgedrückt werden (6 von 86 = 7 %) als durch literale (67 %); dies könnte man als Indiz dafür interpretieren, dass die verblasste Semantik die wertende Verwendung (besser) ermöglicht und damit primär ist.

Einen etwas anderen, aber ebenfalls die evaluative Funktion betreffenden Ansatz verfolge ich derzeit zusammen mit einem korpuslinguistischen Kollegen (vgl. Proseke/Zeschel 2023). Wir setzen dafür bei Traugotts (1988) Konzept der Subjektivität an und kodieren diese anhand sprachlicher Merkmale, z. B. der Kombination von deontischem Modalverb und Negation (*ich kann nicht kommen und sagen*, vgl. a. Beispiel 2) und expliziter Wertungen, die dem pseudokoordinierten Satz vorausgehen oder folgen, z. B. *cool* oder *ätzend* (vgl. 9 und 10). Die Kodierung unterscheidet nur in subjektiv und nicht subjektiv, die Indikatoren dafür werden separat kodiert. Der bisherige Befund ist, dass die pseudokoordinierten Bewegungsverben *kommen* und *gehen* dann deutlich häufiger subjektiv verwendet werden, wenn sie mit einer Verbpartikel wie *an-*, (*da*)*her-* oder *hin-* kombiniert sind. Die Auswertung zeigt nebenbei auch, dass keineswegs mehrheitlich eine (negative) Wertung oder ein konfliktärer Kontext vorliegen muss, wenn Pseudokoordination verwendet wird. Die laufende Untersuchung ist keine auf medial mündlichen Daten basierende. Da die in FOLK vorkommenden Belege für Pseudokoordination (mit allen gängigen V1) nicht ausreichen, um statistisch valide Partikelverb- und Nicht-Partikelverb-Varianten zu vergleichen, greifen wir auf Daten aus DECOW16B (Schäfer/Bildhauer 2012) zurück. Dieses Korpus enthält in großen Teilen relativ spontansprachliche Internetkommentare, die teilweise als konzeptionell mündlich kategorisiert werden können. Damit eignet sich das Korpus zum ›Auffüllen‹ von Datenlücken in FOLK und als Vergleichsbasis sowohl in Richtung medial mündlicher Korpora als auch traditioneller Schriftkorpora (vgl. auch Zeschel/Bildhauer/Weber 2023). Wo FOLK zu wenig Belege für z. B. pseudokoordiniertes *hergehen* enthält, sind die Belege in DECOW so zahlreich, dass eine Stichprobe genommen werden muss, die sich auf (nahezu)

¹⁸ Aus Platzgründen kann hier nicht näher auf die Kodierung der Bedeutung eingegangen werden.



direkt adjazente Fälle beschränkt. Dabei eröffnet sich eine Menge ganz anderer Kodier- und Auswertungsreflexionen, z.B. zur Vergleichbarkeit mit den mündlichen Daten, die eine größere Wortstellungsvariation enthalten.

6 Schlussbemerkungen

Die beiden Vergleiche meiner PK-Auswertungen mit Ergebnissen aus der Literatur zu PK in anderen Sprachen haben gezeigt, dass Studien zu demselben Thema, denen unterschiedliche Kodierweisen zugrunde liegen, eher als Bausteine in einem übergeordneten Erkenntnisbauwerk gelesen werden können denn als Replikationen, die Ergebnisse unmittelbar stützen oder widerlegen. Sowohl was die prosodische Gestaltung als auch was die subjektive Funktion der PK und ihren jeweiligen Zusammenhang mit der Semantik angeht, gibt es noch viele offene Fragen. In jedem Fall ist die Kodierung und anschließende Quantifizierung ein zentraler Baustein in aktuellen IL-Arbeiten. Behauptungen aus der Literatur könnten, insbesondere wenn sie als für alle Belege eines Phänomens geltend formuliert wurden, im Prinzip zwar schon anhand einzelner Belege falsifiziert werden, ein Ablehnen anhand von statistischen Generalisierungen ist aber methodologisch deutlich differenzierter. Diese Generalisierungen zu erzielen, ist bei Korpusdaten, die Kodierungen erfordern, nie trivial, weist aber bei mündlichen Daten zusätzliche Besonderheiten auf. Neben der Reflexion von Kodierweisen ist es notwendig, stets auch die möglichen, verschiedenen Interpretationen der fremden und eigenen Ergebnisse ausgiebig zu reflektieren.

Wie steht es nun um das Verhältnis dieser Reflexionen zur Hermeneutik? Hermanns (2009) bezeichnet die gesamte Linguistik als Verstehenslehre, weil sie auf einer »praktischen Hermeneutik« gründe, das heißt, in allen linguistischen Subdisziplinen »ist linguistische Erkenntnis ohne eigenes Sprachverstehen von SprachwissenschaftlerInnen gar nicht möglich, die insoweit niemals gänzlich »unbeteiligte Beobachter« sein können (Schermer 1984, S. 60).« (Hermanns 2009, S. 181) Diese Sicht auf das methodische Nutzen des eigenen Sprachverstehens durch die Analytiker:innen hat, wie in Abschnitt 4 thematisiert, die CA zum allgemeinen Prinzip erhoben und damit den Schwerpunkt sogar umgekehrt: Der unbeteiligte Beobachter spielt im Idealfall keine Rolle, in der Realität aktueller IL-Arbeiten ist er zwar präsent und wird aufgewertet, er ist der Analyse aus Teilnehmendenperspektive aber oft nachgeordnet. Trotz dieses Fokus fragt sich, auf welcher Ebene man hier von Hermeneutik sprechen kann. Traditionell ist die Hermeneutik auf Textrezeption ausgerichtet, nicht auf interaktive Verstehenskonstitution (vgl. Schmitz 2021). Dennoch muss unterschieden werden zwischen dem analytischen Nachvollzug der Verstehensleistungen der Interagierenden und einer durch Analytikerwissen angereicherten Interpretation. Nur für ersteren gilt:

Das Gesprochene ist flüchtig – ein späterer Rückgriff auf das Original zur vertiefenden Auslegung ist nicht möglich. Damit wird für das dialogische Verstehen die hermeneutische Unterscheidung zwischen Verstehen und Auslegen/ Interpretieren hinfällig (Deppermann 2008, S. 229).

Durch Transkription wird diese Trennung dann wieder möglich, auch wenn sie durch die methodologischen Prämissen der CA eigentlich nicht gewünscht ist: Es soll gerade keine von der Teilnehmendenperspektive zu unterscheidende Interpretation der wissenschaftlich Untersuchenden gemacht werden. Wie aber bereits thematisiert, ist der Nutzen von Analytikerwissen mittlerweile teilweise anerkannt, und aus meiner eigenen, mit korpuslinguistischen Ansätzen kombinierten IL-Perspektive ist es zudem notwendig, auch sprachwissenschaftlich etablierte Kategorien bei der Analyse zu nutzen, die für die Interagierenden nicht unmittelbar relevant sind. Insofern wäre es denkbar, bei der für die Kategorisierung nötigen Interpretation von Hermeneutik zu sprechen. Die Vielfalt der Hermeneutik-Begriffe (vgl. Hermanns 2009) und die Tradition in monologischer Schriftlichkeit stehen jedoch m. E. einer einfachen Anwendung des Begriffs ohne definitorische Expositionen entgegen. Bei vielen linguistischen Fragestellungen und Kategorisierungen geht es nicht um Verstehen (vgl. auch Schmitz 2021), weil nicht immer die Semantik oder Pragmatik eines Satzes bzw. das Verstehen von Äußerungen interessiert. Es wird also nicht immer aus wissenschaftlicher Perspektive das Gesagte ausgelegt, sondern es geht um ganz andere, rein wissenschaftliche Auslegungen. Es kommt hinzu, dass sowohl die CA/IL als auch gebrauchsbasierte korpuslinguistische Ansätze gerade keine feste Methodologie bzw. Lehre des Auslegens verfolgen, sondern die interpretative Offenheit und Revidierbarkeit betonen. Damit ist natürlich auch noch keine Definition dessen gegeben, was alles unter *Interpretation* oder *Interpretieren* fällt. Für die Zwecke des vorliegenden Artikels soll *Interpretieren* nicht terminologisch verstanden werden, sondern alle Bedeutungen des Wortes umfassen, die in aktuellen Wörterbüchern zu finden sind und damit alltagssprachliche ebenso wie wissenschaftliche Bedeutungen umfassen. Letztere standen bei den Beispielen unter 5.2 im Fokus, also die durch (fachliches) Vorwissen informierte, begründbare Auslegung von Kategorien (vgl. Fußnote 1).

In Bezug auf derartige Interpretation zeigt sich an IL-Arbeiten ebenso wie an allen linguistischen Arbeiten, die über Einzelfallanalysen hinausgehen und mit Kategorien arbeiten, die Verschränkung von interpretativen Schritten und Quantifizierungen: Eine Quantifizierung ist ohne die Bildung von Kategorien nicht möglich. Sowohl die Bildung von Kategorien (durch das Festlegen von Merkmalen) als auch das Zuordnen von Belegen zu Kategorien (durch das Prüfen und Gewichten von Merkmalen) sind interpretative Verfahren. Andererseits sind eben diese Verfahren häufig schon durch mögliche, antizipierte quantitative Generalisierungen geprägt, wenn beispielsweise entschieden wird, welche Merkmale oder Merkmalsausprägungen überhaupt gezählt werden sollen, weil sie intuitiv nicht zu selten oder irrelevant erscheinen. Insofern sind die allermeisten empirischen linguistischen Arbeiten interpretativen Methoden deutlich verhaftet, was – wenn es reflektiert wird – aber gerade als eine Stärke der Linguistik gesehen werden kann, denn eine Quantifikation nicht qualitativ näher überprüfter Kategorien und ihrer statistischen Zusammenhänge kann interpretativ fragwürdige Ergebnisse liefern (vgl. z. B. auch Bubenhofer 2018; Bubenhofer/Scharloth 2015).

Funding Open Access funding enabled and organized by Projekt DEAL.



Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Barth-Weingarten, Dagmar (2016): *Intonation units revisited. Caesuras in talk-in-interaction*. Amsterdam: Benjamins.
- Barth-Weingarten, Dagmar/Couper-Kuhlen, Elizabeth (2011): Action, prosody and emergent constructions: The case of *and*. In: Peter Auer/Stefan Pfänder (Hg.): *Constructions – emerging and emergent*. Berlin: De Gruyter, S. 263–292.
- Bender, Michael (2024): Korpusgestützte Theoriebildung als hermeneutischer Prozess – iterativ-inkrementelle Entwicklung eines Kategoriensystems am Beispiel einer Theorie des Kommentierens. In: *Zeitschrift für Literaturwissenschaft und Linguistik* 54(2)
- Bergmann, Pia (2022): Sprachliche Variation im Gegenwartsdeutschen: Lautliche Realisierungsvarianten im Gespräch. In: Michael Beißwenger/Lothar Lemnitzer/Carolin Müller-Spitzer (Hg.): *Forschen in der Linguistik. Eine Methodeneinführung für das Germanistik-Studium*. Paderborn: Brill Fink, S. 83–102.
- Birkner, Karin/Auer, Peter/Bauer, Angelika/Kotthoff, Helga (2020): *Einführung in die Konversationsanalyse*. Berlin: De Gruyter.
- Bubenhöfer, Noah (2018): Wenn »Linguistik« in »Korpuslinguistik« bedeutungslos wird. Vier Thesen zur Zukunft der Korpuslinguistik. In: Joachim Gessinger/Angelika Redder/Ulrich Schmitz (Hg.): *Korpuslinguistik*. (= Osnabrücker Beiträge zur Sprachtheorie (OBST)). Duisburg: Universitätsverlag Rhein-Ruhr OHG, S. 17–30.
- Bubenhöfer, Noah/Scharloth, Joachim (2015): Maschinelle Textanalyse im Zeichen von Big Data und Data-driven Turn – Überblick und Desiderate. In: *Zeitschrift für Germanistische Linguistik* 43(1), S. 1–26.
- Couper-Kuhlen, Elizabeth/Selting, Margret (2018): *Interactional Linguistics. Studying language in social interaction*. Cambridge: Cambridge University Press.
- Deppermann, Arnulf (2000): Ethnographische Gesprächsanalyse: Zu Nutzen und Notwendigkeit von Ethnographie für die Konversationsanalyse. In: *Gesprächsforschung* 1, S. 96–124.
- Deppermann, Arnulf (2008): Verstehen im Gespräch. In: Heidrun Kämper/Ludwig M. Eichinger (Hg.): *Sprache - Kognition - Kultur. Sprache zwischen mentaler Struktur und kultureller Prägung*. Berlin/New York: De Gruyter, S. 225–261.
- Deppermann, Arnulf (2012): How does ›cognition‹ matter to the analysis of talk-in-interaction? In: *Language Sciences* 34, S. 746–767.
- Deppermann, Arnulf/Reineke, Silke (2023): Zur Verwendung von Metadaten in der interaktionsanalytischen Arbeit mit Korpora – am Beispiel einer Untersuchung anhand des Korpus FOLK. In: Michael Beißwenger/Eva Gredel/Lothar Lemnitzer/Roman Schneider (Hg.): *Korpusgestützte Sprachanalyse. Grundlagen, Anwendungen und Analysen*. Tübingen: Narr Francke Attempto, S. 245–259.
- Deppermann, Arnulf/Proske, Nadine/Zeschel, Arne (Hg.) (2017): *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*. Tübingen: Narr.
- Dingemans, Mark/Enfield, N.J. (2015): Other-initiated repair across languages: towards a typology of conversational structures. In: *Open Linguistics* 1, S. 96–118.
- Dingemans, Mark/Kendrick, Robin/Enfield, N.J. (2016): A coding scheme for other-initiated repair across languages. In: *Open Linguistics* 2, S. 35–46.

- Drew, Paul/Hakulinen, Auli/Heinemann, Trine/Niemi, Jarkko/Rossi, Giovanni (2021): Hendiadys in naturally occurring interactions: A cross-linguistic study of double verb constructions. In: *Journal of Pragmatics* 182, S. 322–347.
- Egbert, Maria (2009): *Der Reparatur-Mechanismus in deutschen Gesprächen*. Mannheim: Verlag für Gesprächsforschung.
- Ehmer, Oliver (2011): *Imagination und Animation. Die Herstellung mentaler Räume durch animierte Rede*. Berlin/New York: De Gruyter.
- Ehmer, Oliver (2022): *Makrokonstruktionen. Komplexe Adverbialstrukturen zwischen lokaler Emergenz und Sedimentierung im gesprochenen Französisch*. Berlin/Boston: De Gruyter.
- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (Hg.) (2017): *Gesprochene Wissenschaftssprache – digital: Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen: Stauffenburg.
- Floyd, Simeon/Rossi, Giovanni/Enfield, N.J. (2020a): A coding scheme for recruitment sequences in interaction. In: Simeon Floyd/Giovanni Rossi/N. J. Enfield (Hg.): *Getting others to do things: A pragmatic typology of recruitments*. Berlin: Language Science Press, S. 25–50. 36.
- Floyd, Simeon/Rossi, Giovanni/Enfield, N.J. (Hg.) (2020b): *Getting others to do things: A pragmatic typology of recruitments*. Berlin: Language Science Press.
- Günthner, Susanne (2000): *Vorwurfsaktivitäten in Alltagsinteraktionen. Grammatische, prosodische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommunikativer Muster und Gattungen*. Tübingen: Niemeyer.
- Günthner, Susanne/Hopper, Paul J. (2010): Zeitlichkeit & sprachliche Strukturen: Pseudoclefts im Englischen und Deutschen. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 11, S. 1–28.
- Haddington, Pentti/Jantunen, Jarmo H./Sivonen, Jari (2011): Language and affect: Go-say and come-say constructions in Finnish. In: *Sky Journal of Linguistics* 24, S. 75–117.
- Helmer, Henrike/Zinken, Jörg (2019): *Das heißt* (»that means«) for formulations and *du meinst* (»you mean«) for repair? Interpretations of prior speakers' turns in German. In: *Research on Language and Social Interaction* 52(2), S. 159–176.
- Hermans, Fritz (2009): Linguistische Hermeneutik. Überlegungen zur überfälligen Einrichtung eines in der Linguistik bislang fehlenden Teilfaches. In: Ekkehard Felder (Hg.): *Sprache*. Heidelberger Jahrbücher 53. Berlin/Heidelberg: Springer, S. 179–214.
- Hopper, Paul J. (2002): Hendiadys and auxiliation in English. In: Joan Bybee/Michael Noonan (Hg.): *Complex sentences in grammar and discourse. Essays in honor of Sandra A. Thompson*. Amsterdam/Philadelphia: Benjamins, S. 145–173.
- Hopper, Paul J. (2008): Emergent serialization in English: Pragmatics and typology. In: Jeff Good (Hg.): *Linguistic universals and language change*. Oxford: Oxford University Press, S. 253–284.
- Imo, Wolfgang/Lanwer, Jens Philipp (2019): *Interaktionale Linguistik. Eine Einführung*. Berlin: Metzler.
- Kaiser, Julia (2016): Reformulierungsindikatoren im gesprochenen Deutsch. Die Benutzung der Ressourcen DGD und FOLK für gesprächsanalytische Zwecke. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 17, S. 196–230.
- Kaiser, Julia/Schmidt, Thomas (2016): Einführung in die Benutzung der Ressourcen DGD und FOLK für gesprächsanalytische Zwecke. Handreichung: Einfache Recherche-Anfragen als Übungsbeispiele. Mannheim: IDS-Publikationsserver. urn:nbn:de:bsz:mh39-55360
- Kendrick, Kobin H./Drew, Paul (2016): Recruitment: offers, requests, and the organization of assistance in interaction. In: *Research on Language and Social Interaction* 49(1), S. 1–19.
- Kornfeld, Laurenz/Küttner, Uwe-A./Zinken, Jörg (2023): Ein Korpus für die vergleichende Interaktionsforschung. Das »Parallel European Corpus of Informal Interaction« (PECII). In: Arnulf Deppermann/Christian Fandrych/Marc Kupietz/Thomas Schmidt (Hg.): *Korpora in der germanistischen Sprachwissenschaft. Mündlich, schriftlich, multimedial*. Berlin: De Gruyter, S. 103–127.
- Küttner, Uwe-A./Kornfeld, Laurenz/Zinken, Jörg (2023): A coding scheme for (dis-)approval-relevant events involving the direct social sanctioning of problematic behavior in informal social interaction. Mannheim: IDS-Verlag. <https://doi.org/10.21248/idsopen.5.2023.8>.
- Lanwer, Jens Philipp (2018a): Was steck in den Daten und was stecken wir hinein? Möglichkeiten und Grenzen der gebrauchsbasierten Rekonstruktion sprachlicher Strukturen. In: Joachim Gessinger/Angelika Redder/Ulrich Schmitz (Hg.): *Korpuslinguistik*. (= Osnabrücker Beiträge zur Sprachtheorie (OBST)). Duisburg: Universitätsverlag Rhein-Ruhr OHG, S. 219–238.
- Lanwer, Jens Philipp (2018b): Grammatikalität und Rekurrenz. Zur Rolle statistischer Verfahren im Rahmen einer »rekonstruktiven« Linguistik. In: Georg Albert/Sabine Diao-Klaeger (Hg.): *Mündlicher Sprachgebrauch zwischen Normorientierung und pragmatischen Spielräumen*. Tübingen: Stauffenburg, S. 232–53.

- Lanwer, Jens Philipp (2020): Appositive Syntax oder appositive Prosodie? In: Wolfgang Imo/Jens Philipp Lanwer (Hg.): *Prosodie und Konstruktionsgrammatik*. Berlin: De Gruyter, S. 233–281.
- Müller, Marcus (2024): Einsam oder gemeinsam? Verstehen und Erklären in der digitalen Linguistik. In: *Zeitschrift für Literaturwissenschaft und Linguistik* 54(2).
- Mundwiler, Vera/Kreuz, Judith/Müller-Feldmeth, Daniel/Luginbühl, Martin/Hauser, Stefan (2019): Quantitative und qualitative Zugänge in der Gesprächsforschung – Methodologische Betrachtungen am Beispiel einer Studie zu argumentativen Gruppendiskussionen. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 20, S. 323–383.
- Proske, Nadine (2017): Perspektivierung von Handlungen und Zuschreibung von Intentionalität durch pseudokoordinierte kommen. In: Arnulf Deppermann/Nadine Proske/Arne Zeschel (Hg.): *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*. Tübingen: Narr, S. 177–247.
- Proske, Nadine (2019): Emergent pseudo-coordination in spoken German. A corpus-based exploration. In: *Yearbook of the German Cognitive Linguistics Association (GCLA)* 7, S. 115–136.
- Proske, Nadine (2023): Pseudo-coordinated *sitzen* and *stehen* in spoken German – a case of emergent progressive aspect? In: *Journal of Germanic Linguistics* 35(4), S. 447–486.
- Proske, Nadine/Deppermann, Arnulf (2020): Right-dislocated complement clauses in German talk-in-interaction: (Re-)specifying propositional referents of the demonstrative pronoun *das*. In: Yael Maschler/Simona Pekarek Doehler/Jan Lindström/Leelo Keevallik (Hg.): *Emergent syntax for conversation. Clausal patterns and the organization of action*. Amsterdam: Benjamins, S. 275–302.
- Proske, Nadine/Zeschel, Arne (2023): Pseudo-coordination in German. Vortrag auf der »16th International Cognitive Linguistics Conference«, Düsseldorf, 09.08.2023.
- Reineke, Silke/Schmidt, Thomas (2022): Das Archiv für Gesprochenes Deutsch und das Forschungs- und Lehrkorpus für Gesprochenes Deutsch. Audiovisuelle Dokumentation von Sprachgebrauch in Gesellschaft und Politik. In: Heidrun Kämper/Albrecht Plewnia (Hg.): *Sprache in Politik und Gesellschaft. Perspektiven und Zugänge*. Berlin: De Gruyter, S. 323–330.
- Reineke, Silke/Deppermann, Arnulf/Schmidt, Thomas (2023): Das Forschungs- und Lehrkorpus für Gesprochenes Deutsch (FOLK). Zum Nutzen eines großen annotierten Korpus gesprochener Sprache für interaktionslinguistische Fragestellungen. In: Arnulf Deppermann/Christian Fandrych/Marc Kupietz/Thomas Schmidt (Hg.): *Korpora in der germanistischen Sprachwissenschaft. Mündlich, schriftlich, multimedial*. Berlin: De Gruyter, S. 71–102.
- Schäfer, Roland/Bildhauer, Felix (2012): Building large corpora from the web using a new efficient tool chain. In: Nicoletta Calzolari/Khalid Choukri/Thierry Declerck/Mehmet Ugur Dogan/Bente Maegaard/Joseph Mariani/Jan Odijk/Stelios Piperidis (Hg.): *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul: ELRA, S. 486–493.
- Schegloff, Emanuel A. (1993): Reflections on quantification in the study of conversation. In: *Research on Language and Social Interaction* 26(1), S. 99–128.
- Scherner, Maximilian (1984): Sprache als Text. Ansätze zu einer sprachwissenschaftlich begründeten Theorie des Textverstehens. In: Gerhard Stickel (Hg.): *Stilfragen*. Tübingen: Niemeyer.
- Schmidt, Thomas (2017a): DGD – Die Datenbank für Gesprochenes Deutsch. Mündliche Korpora am Institut für Deutsche Sprache (IDS) in Mannheim. In: *Zeitschrift für Germanistische Linguistik* 45(3), S. 451–463.
- Schmidt, Thomas (2017b): Construction and Dissemination of a Corpus of Spoken Interaction – Tools and Workflows in the FOLK project. In: *Journal for Language Technology and Computational Linguistics* 31(1), S. 127–154.
- Schmidt, Thomas (2018): Gesprächskorpora. Aktuelle Herausforderungen für einen besonderen Korpus-typ. In: Marc Kupietz/Thomas Schmidt (Hg.): *Korpuslinguistik*. Berlin: De Gruyter, S. 209–230.
- Schmidt, Thomas (2022): Korpora gesprochener Sprache. In: Michael Beißwenger/Lothar Lemnitzer/Carolin Müller-Spitzer (Hg.): *Forschen in der Linguistik. Eine Methodeneinführung für das Germanistik-Studium*. Paderborn: Brill Fink, S. 421–430.
- Schmitz, Ulrich (2021): Sprachwissenschaft und Verstehenskunst: Zwei ungleiche Schwestern. In: *Zeitschrift für Literaturwissenschaft und Linguistik* 51, S. 685–693.
- Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner, Karin/Couper-Kuhlen, Elizabeth/Deppermann, Arnulf/Gilles, Peter/Günthner, Susanne/Hartung, Martin/Kern, Friederike/Mertzluff, Christine/Meyer, Christian/Morek, Miriam/Oberzaucher, Frank/Peters, Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock, Anja/Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, S. 353–402.

- Sidnell, Jack/Stivers, Tanya (Hg.) (2013): *The handbook of Conversation Analysis*. Malden: Wiley-Blackwell.
- Stensig, Jakob/Heinemann, Trine (2015): Opening up codings? In: *Research on Language and Social Interaction* 48(1), S. 20–25.
- Stefanowitsch, Anatol (2000): The English GO-(PRT)-AND-VERB construction. In: *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society*, S. 259–270.
- Stift, Ulf-Michael/Schmidt, Thomas (2014): Mündliche Korpora am IDS: Vom Deutschen Spracharchiv zur Datenbank für Gesprochenes Deutsch. In: Institut für Deutsche Sprache (Hg.): *Ansichten und Einsichten. 50 Jahre Institut für Deutsche Sprache*. Mannheim: Institut für Deutsche Sprache, S. 360–375.
- Stivers, Tanya (2015): Coding social interaction: a heretical approach in Conversation Analysis? In: *Research on Language and Social Interaction* 48(1), S. 1–19.
- Stivers, Tanya/Enfield, N.J. (2010): A coding scheme for question-response sequences in conversation. In: *Journal of Pragmatics* 42, S. 2620–2626.
- Traugott, Elizabeth Closs (1988): Pragmatic strengthening and grammaticalization. In: *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, S. 406–416.
- Westpfahl, Swantje (2020): *POS-Tagging für Transkripte gesprochener Sprache. Entwicklung einer automatisierten Wortarten-Annotation am Beispiel des Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK)*. Tübingen: Narr.
- Westpfahl, Swantje/Proske, Nadine/Hobich, Melanie/Borlinghaus, Anton/Strub, Hanna (2018): Guideline: Syntactic annotation and segmentation in the SegCor Project. Mannheim: IDS-Publikationsserver. <https://doi.org/10.14618/ids-pub-8919>.
- Westpfahl, Swantje/Schmidt, Thomas/Borlinghaus, Anton/Strub, Hanna (2019): Guideline: Syntaktische Segmentierung in FOLKER. Mannheim: IDS-Publikationsserver. <https://doi.org/10.14618/ids-pub-8762>.
- Zeschel, Arne/Bildhauer, Felix/Weber, Thilo (2023): Ongoing change and where to track it. Incipient grammaticalization and the German *am*-progressive. Vortrag auf der »16th International Cognitive Linguistics Conference«, Düsseldorf, 09.08.2023.
- Zima, Elisabeth (2013): *Kognition in der Interaktion. Eine kognitiv-linguistische und gesprächsanalytische Studie dialogischer Resonanz in österreichischen Parlamentsdebatten*. Heidelberg: Universitätsverlag Winter.
- Zinken, Jörg/Deppermann, Arnulf (2017): A cline of visible commitment in the situated design of imperative turns. Evidence from German and Polish. In: Marja-Leena Sorjonen/Liisa Raevaara/ Elizabeth Couper-Kuhlen (Hg.): *Imperative Turns at Talk. The design of directives in action*. Amsterdam: Benjamins, S. 27–63.
- Zinken, Jörg/Küttner, Uwe-A. (2022): Offering an interpretation of prior talk in everyday interaction: A semantic map approach. In: *Discourse Processes* 59(4), S. 298–325.

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.

