

Malte Belz und Carolin Odebrecht\*

# Abschnittsweise Analyse sprachlicher Flüssigkeit in der Lernaltersprache: Das Ganze ist weniger informativ als seine Teile

## Section-by-section analysis of linguistic fluency in learner language: The whole is less informative than its parts

<https://doi.org/10.1515/zgl-2022-2051>

**Abstract:** In this corpus-based study we explore three measurements of L2 fluency – articulation rate, filler particles, and pauses –, both within and between two registers of spontaneous dialogues spoken by Polish learners of German. The measurements are assessed both *in toto* (as calculated over the whole dialogue) and *in parte* (as calculated for specific sections). The sections are identified on a quantitative tier that divides the dialogue into four parts, and qualitatively on two linguistically-informed tiers, comprising sections based on dialogue move and task. We challenge the assessment of fluency as an average measurement over the entire dialogue, showing that a sectionwise analysis offers a better understanding of similarities and differences both within and between the two registers.

- 1 Einführung
- 2 L2-Flüssigkeit und Register
- 2.1 Flüssigkeitsmaße
- 2.2 Register und gesprochensprachliche Phänomene
- 3 Methode
- 3.1 Korpus
- 3.2 Annotation
- 3.3 Auswertung
- 4 Ergebnisse
- 4.1 Quantitative Abschnitte
- 4.2 Dialogzüge

---

**\*Kontaktpersonen:** Dr. Malte Belz: Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik, Unter den Linden 6, D-10099 Berlin, E-Mail: malte.belz@hu-berlin.de  
Dr. Carolin Odebrecht: Humboldt-Universität zu Berlin, Sprach- und literaturwissenschaftliche Fakultät & Computer- und Medienservice, Unter den Linden 6, D-10099 Berlin, E-Mail: carolin.odebrecht@hu-berlin.de

- 4.3 Aufgabenabschnitte der Diapix-Aufgabe
- 5 Diskussion  
Literatur

## 1 Einführung

Sprachliche Flüssigkeit ist ein Konstrukt. Fillmore (1979, S. 93) beschreibt sprachliche Flüssigkeit als „the ability to talk at length with few pauses, the ability to fill time with talk [...], the ability to talk in coherent [...] sentences [...], the ability to have appropriate things to say in a wide range of contexts [...] and the ability to be creative and imaginative in [...] language use“. Für den Zweitspracherwerb (L2-Erwerb) ist jedes Merkmal dieser ‚mündlichen Sprachfertigkeiten‘ für sich genommen eine Herausforderung. Dieser Artikel lässt Flüssigkeit im ‚weiten‘ Sinne („fließende“ Sprachverwendung) außen vor und untersucht Flüssigkeit im ‚engen‘ Sinne (Lennon, 1990), wobei mögliche Merkmale über die Fillmoresche Definition weniger Pausen hinausgehen.<sup>1</sup> Die Betrachtung von Flüssigkeit im engen Sinne ermöglicht die Definition von isolierbaren Komponenten der mündlichen Sprachverwendung wie beispielsweise Artikulationsgeschwindigkeit oder Pausendauer. In dieser Studie verwenden wir diesen engen Blick auf sprachliche Flüssigkeit, der es erlaubt, Flüssigkeit mithilfe von Flüssigkeitsmaßen zu operationalisieren und auszuwerten. Diese Operationalisierung von Flüssigkeit kommt ohne muttersprachliche Vergleichsgruppe aus. Stattdessen vergleichen wir drei Flüssigkeitsmaße innerhalb und zwischen zwei Registern. Sprachgebrauch ist registersensitiv (Biber & Conrad, 2009; Labov, 1997). Das heißt, die Frequenz und Distribution von linguistischem Material hängt stark von situativen und funktionalen Umständen ab. Linguistisch wird typischerweise erforscht, welche konstituierenden Gemeinsamkeiten oder distinktiven Unterschiede zwischen Registern hinsichtlich der Distribution linguistischer Merkmale existieren (vgl. z. B. Biber, 2009). Unser Beitrag weitet den Untersuchungsgegenstand auf Phänomene gesprochener Sprache aus und untersucht, ob Register in Dialogen auch hinsichtlich der Distribution von Flüssigkeitsmerkmalen unterschieden werden können. Wir stellen explorativ die Hypothese auf, dass diese Merkmale ebenfalls funktional-situativ von Sprecher\*innen eingesetzt werden (können) und schließen damit an erste frühe Studien an, die Unterschiede zwischen Selbstbeschreibungen, Cartoonbeschreibungen und der Erklärung von Sprichwörtern aufgezeigt

---

<sup>1</sup> In neueren Modellierungen wird auch vom ‚sehr engen‘ Sinne gesprochen (Tavakoli & Hunter, 2018).

haben (Lay & Paivio, 1969). Lerner\*innen zeigen – zumindest schriftsprachlich – weniger Registerunterschiede im Gebrauch als Muttersprachler\*innen (vgl. für Englisch als L2 Gilquin & Paquot, 2008). Da unser Beitrag besonders auf einen methodologisch-explorativen Erkenntnisgewinn abzielt, der über erste Beschreibungsmodelle nicht hinausgeht, konzentrieren wir uns bei der Untersuchung der Flüssigkeitsmaße auf Lernerdaten in zwei Registern. Eine umfassende theoretische Einbettung und Diskussion könnte später zum Beispiel im Rahmen der systematischen funktionalen Linguistik (Halliday, 2014) erfolgen – dann auch mit kontrastiven Vergleichen. Für Flüssigkeitsmerkmale wurde diese Hypothese nach unserer Kenntnis allerdings weder für L1 noch L2 systematisch exploriert.

Zusätzlich erweitern wir diesen Ansatz paradigmatisch mit der Betrachtung von Abschnitten innerhalb der Dialoge. Neben der Zuordnung zu einem Register kann ein Dialog während seines Verlaufs verschiedene Phasen oder Abschnitte enthalten, die sich wiederum untereinander bezüglich der Frequenz und Distribution von linguistischen Merkmalen unterscheiden. So ist es nicht zwingend, dass sprachliche Flüssigkeit beziehungsweise deren Ausprägung über den Dialog hinweg gleich verteilt ist – sie kann sich auch in bestimmten Abschnitten verstärken oder reduzieren, zum Beispiel, wenn Füllpartikeln zusammen mit Pausen, Reparaturen oder Prolongationen auftreten (vgl. Belz et al., 2017; Betz, 2020; Clark & Fox Tree, 2002). Ähnliche Ansätze abschnittsweiser Analysen verfolgen auch Kerz et al. (2020) bei der Untersuchung des Schriftspracherwerbs, Schnelle (2018) für die historische Linguistik oder Reagan et al. (2016) in der Literaturwissenschaft (obwohl bei textbasierten Daten strenggenommen eine zeitliche Abfolge fehlt). Studien, die ein Werk ausschließlich in seiner Gesamtheit betrachten (z. B. in einem Maß X pro Wort), können so schwerlich Clustereffekte zwischen Abschnitten oder (Sub-)Registern innerhalb eines Werks identifizieren.

In der Kombination beider Analyseansätze untersuchen wir korpusbasiert drei Maße für sprachliche Flüssigkeit bei Deutschlerner\*innen in spontansprachlichen Dialogen *in toto* (über den gesamten Dialog hinweg ermittelt) und *in parte* (über quantitativ und qualitativ motivierte und explizit modellierte Abschnitte hinweg ermittelt): Artikulationsgeschwindigkeit, Füllpartikeln und Pausenanzahl. Dazu verwenden wir das Wrocław Dialogue Corpus v.2 (Wesolek & Belz, 2021). In WroDiaCo unterhalten sich polnische Deutschlerner\*innen in einem themenorientierten, eher freien und einer aufgabenorientierten Kommunikationssituation (vgl. Abschnitt 3), illustriert mit dem nachfolgenden qualitativ (ohne Zeitalignierung) wiedergegebenen Dialogauszug (Paar i, Diapix<sup>2</sup>, Sek. 786).

---

2 Eine Diapix-Aufgabe besteht aus zwei fast gleichen Zeichnungen mit zehn Unterschieden, welche beide Gesprächspartner\*innen finden müssen, ohne die jeweils andere Zeichnung zu sehen. Vergleiche dazu auch Hirschmann und Schmidt (in diesem Heft) und Trouvain (in diesem Heft).

- (1) **SP1** und was ist noch mit dem Schild <p> bei dir <p> ist noch <p>  
**SP2** bei mir ist ein Unterschrift also äh das heißt <p> äh ein Text  
 Fleischer  
**SP1** ha <p> das gibt es nicht hier <p>  
**SP2** <usb> <p> okay <p>  
**SP1** dann vierter <p>  
**SP2** also vielleicht haben wir vier <p> Unterschiede

Der Auszug in Beispiel 1 enthält die diplomatische Transliteration<sup>3</sup> der Sprecher\*innen SP1 und SP2 (Ebene *ortflue*) und zeigt typische spontansprachliche Merkmale wie Pausen (<p>) und Füllpartikeln (*äh*), während die Sprecher\*innen versuchen, zehn Unterschiede auf den ihnen vorliegenden Bildern zu finden.<sup>4</sup> Dies machen sie am Ende dieses vierten Unterschieds explizit. Diese aufgabeninduzierten Abschnitte bis zum Finden des nächsten Unterschiedes stellt eine der Abschnittsebenen dar, die wir in dieser Studie analysieren. Diese und weitere *In parte*-Analysen können aufgrund der linguistisch-funktional motivierten qualitativen Annotationsabschnitte nur statisch miteinander verglichen werden (für dynamische Vergleiche s. Kerz et al., 2020).

Zusammenfassend verfolgen wir eine methodologisch-explorative (I) und eine inhaltlich-explorative Forschungsfrage (II):

**Forschungsfrage I:** Können abschnittsweise (*in parte*) ermittelte Flüssigkeitsmaße andere oder komplementäre Aussagen über einen Dialog treffen als *In toto*-Maße?

**Forschungsfrage II:** Können Flüssigkeitsmaße im Vergleich von *in parte* vs. *in toto* als Indikator herangezogen werden, um Register zu unterscheiden?

Diese Forschungsfragen werden wir korpusbasiert beantworten. Wir argumentieren dafür, dass Oberflächenmerkmale *in toto* keine ausreichenden Aussagen über sprachliche Flüssigkeit treffen können und schlagen einen modellgestützten Begriff von Flüssigkeitsmaßen vor, der Register und Sprecher\*inneneigenschaften genauer abbilden kann.

<sup>3</sup> Die diplomatische Transliteration orientiert sich sowohl an der Orthographie (ohne Satzzeichen) als auch an der phonetischen Realisation (bspw. *isses* für *ist es*).

<sup>4</sup> Extralinguistische Geräusche sind mit <usb> annotiert.

## 2 L2-Flüssigkeit und Register

Sprachliche Flüssigkeit wird je nach Fachgebiet unterschiedlich definiert. Sie lässt sich mindestens in kognitive Flüssigkeit, wahrgenommene Flüssigkeit und Äußerungsflüssigkeit gliedern (Segalowitz, 2016). In dieser Studie beschränken wir uns auf die zeitlichen Aspekte im akustischen Signal („temporal fluency“) und verwenden statt Äußerungsflüssigkeit den neutraleren Begriff der Oberflächenflüssigkeit („surface fluency“) (Lickley, 2015), da diese unabhängig von den einzelnen, gesprächslinguistisch definierten *Äußerungen* der Sprecher\*innen ist. Für die Forschung über sprachliche Flüssigkeit im Lernerkontext wurde lange angenommen, dass man von einer zielsprachlichen Flüssigkeit nur dann sprechen kann, wenn bestimmte Parameter des Redeflusses denen von Muttersprachler\*innen ähneln und zum Beispiel nicht von stillen Pausen oder Häsitationen unterbrochen wird (vgl. Lennon, 1990, S. 390). Diese Annahme wird mittlerweile nicht mehr getroffen – vielmehr wird in kontrastiven Studien untersucht, ob und wie Lerner\*innen ein zielsprachliches Flüssigkeitsniveau adaptieren (Belz et al., 2017; Gut & Fuchs, 2017; Guz, 2015). Methodologisch sind dennoch zwei Punkte zu hinterfragen.

Zum einen wird sprachliche Oberflächenflüssigkeit typischerweise über die gesamte sprachliche Artikulationszeit je Sprecher\*in berechnet. Beispielsweise werden Füllpartikeln pro Minute oder Artikulationsgeschwindigkeit *in toto*, also im Mittel des gesamten Dialogs berechnet. Dies ist problematisch, da sich Dialoge typischerweise durch Redewechsel und damit durch eine möglicherweise ungleich verteilte Gesamtredezeit je Sprecher\*in auszeichnen.

Zum anderen ist anzunehmen, dass auch für Flüssigkeitsmerkmale registerspezifische Unterschiede bestehen, da Sprecher\*innen verschiedene linguistische Mittel verwenden, um in der kommunikativen Situation des jeweiligen Registers erfolgreich funktional zu handeln (Biber & Conrad, 2009). Damit einher geht, dass auch immer die sprecherspezifische Verwendung von Merkmalen kontrolliert werden muss (vgl. z. B. Klug, 2013; Künzel, 1987; McDougall & Duckworth, 2018) und dass registerspezifische Muster inkonsistent sein können (vgl. Abschnitt 2.2). Zusätzlich besteht die Möglichkeit, dass auch innerhalb eines Dialogs verschiedene Register verwendet werden oder ein Register starke Variation hinsichtlich der Distribution seiner Merkmale aufweist. Die Betrachtung des Verlaufs oder einzelner Abschnitte innerhalb der Dialoge erscheint damit zwingend erforderlich.

### 2.1 Flüssigkeitsmaße

Die Oberflächenflüssigkeit wurde in jüngerer Zeit mithilfe dreier Dimensionen operationalisiert: *speed fluency* – das Tempo, mit dem Sprache produziert wird;

*breakdown fluency* – Unterbrechungen im Redefluss, meistens Pausen; und *repair fluency* – das sind Phänomene syntaktischer Art wie Falschstarts, Reparaturen oder Wiederholungen (vgl. Derwing, 2017; Skehan, 2003; Tavakoli & Skehan, 2005). Als vierte Dimension nehmen wir hier Füllpartikeln an, die manchmal unter *breakdown fluency* subsumiert werden (vgl. Lintunen et al., 2020, S. 5). Im Folgenden werden die Flüssigkeitsmaße für die drei Dimensionen Tempo, Redeflussunterbrechung und Füllpartikeln diskutiert (für einen exhaustiven Überblick vgl. Lintunen et al., 2020). *Repair fluency* wird aufgrund ihrer niedrigen Frequenz und ihrer Kovarianz mit den anderen Flüssigkeitsmaßen ausgeblendet (vgl. Belz et al., 2017).

Die Definition von Flüssigkeitsmaßen ist ein expliziter Schritt, der sich nicht automatisch aus dem sprachlichen Material ableitet oder gar konstituierend für dieses Material ist. Moisl formuliert diese Diskrepanz so:

Data is ontologically different from the world. The world is as it is; data is an interpretation of it for the purpose of scientific study. The weather is not the meteorologist's data – measurements of such things as air temperature are. A text corpus is not the linguist's data – measurements of such things as average sentence length are. (Moisl, 2009, S. 876)

Analog dazu ist Flüssigkeit als abstraktes Konstrukt nicht das untersuchte Datum, sondern ihre Ausprägung als Flüssigkeitsmaß. Die Herausforderung liegt in der Definition und Motivation zweckgerechter Maße für die Bestimmung von L2-Flüssigkeit unter Berücksichtigung der Registereigenschaften für die spontane (L2-)Sprache.

Tempo (*speed fluency*) umfasst Sprechgeschwindigkeit und Artikulationsgeschwindigkeit. Im Maße der Sprechgeschwindigkeit ist das gesamte linguistische und nicht-linguistische Sprachmaterial inklusive aller Pausen und Füllpartikeln enthalten, während Artikulationsgeschwindigkeit das linguistische Sprachmaterial inklusive aller Füllpartikeln, aber ohne Pausen enthält (Laver, 1994). Gemessen wird typischerweise in Silben je Sekunde. In dieser Studie betrachten wir nur die Artikulationsgeschwindigkeit, da diese nicht wie die Sprechgeschwindigkeit von der Pausenlänge als kovariierende Variable beeinflusst wird. Ein drittes Maß für Tempo ist die Lauflänge (*length of run*). Da dieses jedoch hochgradig unterschiedlich definiert wird und sich je nach Studie auf Intervalle zwischen unterschiedlich langen Grenzmarkierungen (400 ms Pause oder Füllpartikel bei Derwing et al., 2004, 3000 ms Pause bei Kormos und Dénes, 2004) oder auf prosodische Einheiten wie Intonationsphrasen (Gut & Fuchs, 2017) bezieht, betrachten wir es hier nicht weiter.

Redeflussunterbrechungen (*breakdown fluency*) können *sensu stricto* nur Pausen und nicht Füllpartikeln sein, da letztere keine Pausen im Sprachsig-

nal darstellen, sondern meistens aus Vokalen oder Vokal-Nasal-Verbindungen bestehen (*äh, ähm*). Andere Phänomene gesprochener Sprache ohne solche Sprachlaute, wie Atmungspausen und auch phonetisch noch näher zu definierende Artikulationsgeräusche vor, nach und während Atmungspausen können hingegen als (Teil einer) Pause klassifiziert werden (Belz & Trouvain, 2019). Flüssigkeitsmaße wie Pausenanzahl und mittlere Pausenlänge sind demnach stark von der verwendeten Kategorisierung einer Studie abhängig, aber auch von der Definition ihrer Mindestlänge. Als untere Schwelle für die Pausendauer werden sehr verschiedene Maße herangezogen – in unserem Beitrag definieren wir eine Schwelle von 50 ms, die artikulatorisch noch sinnvoll ist und die keine Verzerrung der Mittelwerte nach oben zur Folge hat (vgl. Włodarczak & Wagner, 2013).

Füllpartikeln werden orthographisch typischerweise mit *äh, ähm, hm* repräsentiert und häufig als „gefüllte Pausen“ bezeichnet. Hier verwenden wir die Bezeichnung *Füllpartikeln*, da der tradierte Begriff der „gefüllten Pause“ aus einer oberflächenorientierten linguistischen Perspektive keine adäquate Beschreibung darstellt. Trotz ihrer großen phonetischen Formvariabilität unterliegen sie allgemeinen prosodischen Prinzipien (Belz, 2021). Auch stark glottalisierte Formen sind möglich (Belz, 2017). Obwohl Füllpartikeln als nicht-verbale Vokalisierungen beschrieben werden (Trouvain & Truong, 2012), ist ihre Oberflächenform eher „wortartiger“ als andere nicht-verbale Vokalisierungen wie Lachen, Räuspern oder Atmen. Füllpartikeln werden in Lernaltersprachen überproportional häufig verwendet (vgl. Götz, 2013; Lennon, 1990). Götz (2019) zeigt, dass ihre Frequenz mit steigendem Fähigkeitsniveau abnimmt und regt daher an, Füllpartikeln im Lernerkontext nicht als Flüssigkeits- sondern als Disfluenzphänomen zu beschreiben. Baker-Smemoe et al. (2014) zählen Füllpartikeln hingegen zu Reparaturflüssigkeit, Kahng (2014) ordnet sie unter Pausenmaße. Trotz oder gerade aufgrund ihrer unterschiedlichen theoretischen Einordnung zählen Füllpartikeln zu den am häufigsten herangezogenen Flüssigkeitsmaßen. Typischerweise wird die absolute Anzahl von Füllpartikeln und ihr relatives Vorkommen je 100 Wörter oder Token<sup>5</sup> in Lernaltersprachen verwendet.

Die vorgestellten Flüssigkeitsmaße werden zur Beantwortung verschiedener Forschungsfragen im Erst- (L1) und Zweit- oder Fremdspracherwerb (L2) verwendet und in kontrastiven Studien miteinander verglichen (Belz et al., 2017; Derwing et al., 2009; Götz, 2013; Guz, 2015; Kahng, 2014; Tavakoli, 2010; Trofimovich & Baker, 2006; van Hest et al., 1997). In longitudinalen Studien wird untersucht, wie die Steigerung des Fähigkeitsniveaus in der L2 mit Maßen der Oberflächenflüssigkeit korreliert (Baker-Smemoe et al., 2014; Derwing et al., 2009; Towell et

---

5 Für eine Diskussion des Tokenbegriffs vgl. Odebrecht et al. (2017).

al., 1996). Alle diese Studien verwenden typischerweise In-toto-Maße zur Bestimmung von Flüssigkeit. Das folgende Kapitel geht auf die registerspezifische und abschnittsweise Betrachtung von Flüssigkeit ein.

## 2.2 Register und gesprochensprachliche Phänomene

Register sind dadurch gekennzeichnet, dass Sprecher\*innen den funktionalen Anforderungen, die sich aus der jeweiligen Äußerungssituation ergeben, sprachlich mit der Verwendung von verschiedenen Mitteln gerecht werden. Biber und Conrad (2009) fassen diese Perspektive wie folgt zusammen:

The underlying assumption of the register perspective is that core linguistic features like pronouns and verbs are functional, and, as a result, particular features are commonly used in association with the communicative purposes and situational context of texts (Biber & Conrad, 2009, S. 2)

Auch sprachliche Flüssigkeit kann funktional sein – im obigen Sinne ein *core linguistic feature*. Daher müssen wir davon ausgehen, dass über sprachliche Flüssigkeit Register unterschieden werden können. Bisherige L1-Studien über gesprochensprachliche Phänomene betrachten die sprachlichen Äußerungen meist *in toto* und vergleichen die Distribution eines Merkmals in verschiedenen Registern. So stellen Lay und Paivio (1969, S. 31) fest, dass die Anzahl an Füllpartikeln mit zunehmender Schwierigkeit einer Aufgabe signifikant ansteigt, von einer Selbstbeschreibung über eine Cartoonbeschreibung hin zu einer Sprichwortaufgabe. Aufgabenunterschiede wie diese können nach unserer Auffassung *a priori* als Registerunterschiede angesehen werden, da verschiedene Aufgabenziele intra-individuell verschiedene funktionale Unterschiede bewirken können – für andere Aufgabenbegriffe wie beispielsweise im Sinne des Task-based Language Teaching vgl. Karges et al., in diesem Heft). Weiterhin werden Register hauptsächlich im L1-Kontext erforscht (für monologisches vs. dialogisches Englisch vgl. Tavakoli, 2016; für Universitätsvorlesungen vs. spontane Dialoge im Portugiesischen vgl. Moniz et al., 2014; für höhere Flüssigkeitsgrade bei Gesprächen über bekannte Informationen und für höhere Flüssigkeit in Mono- als in Dialogen vgl. Foster & Skehan, 1996; Skehan, 2003; für unterschiedliche Füllpartikelraten in Bildbeschreibungsaufgaben vgl. Bortfeld et al., 2001). Das Register kann also Flüssigkeit beeinflussen.

In-parte-Betrachtungen innerhalb eines Registers sind bisher für Dialogstrukturabschnitte betrachtet worden. So verwenden englische Muttersprachler\*innen in kollaborativen Aufgaben Füllpartikeln in Antworten häufiger in



Erzählungen (Lickley, 2001). Auch Antworten in freien Dialogen deutscher Muttersprachler\*innen zeigen signifikant höhere Füllpartikelraten (2,7–3,3 %) als Fragen und Erzählungen (1,1–1,7 %) (Belz, 2021).

Registerspezifische Untersuchungen von L2-Flüssigkeit sind bisher (soweit wir wissen) nur wenige und nur *in toto* durchgeführt worden. Cucchiari et al. (2002) vergleichen Lesesprache und Spontansprache von Niederländischler\*innen. Derwing et al. (2004) vergleichen Flüssigkeitsmaße von Englischler\*innen mit Mandarin als Muttersprache über drei Register hinweg: Bildbeschreibungen, Monologe und Konversationen. Analysiert werden dabei jedoch nur Samples – jeweils ein 30-sekündiger Abschnitt zu Beginn der Bildbeschreibung und des Monologs sowie ein 90-sekündiger Abschnitt zu Beginn des Dialogs. Dieses Sampling aus Daten birgt jedoch die Gefahr, dass diese ersten Abschnitte eines Dialoges strukturell anders geprägt sind als spätere Abschnitte im Dialog und somit nicht stellvertretend für den ganzen Dialog stehen können. Wir zeigen im Ergebnisteil, dass z. B. für freie Dialoge im ersten Viertel eine erhöhte Artikulationsgeschwindigkeit zu beobachten ist.

Die Register in unserer Studie sind gesprochensprachliche Dialoge, die sich durch ihr Kommunikationsziel unterscheiden. Im ersten Register unterhalten sich die Versuchspersonen in einem eher freien, themenorientierten Dialog über ihre Freizeit. Im zweiten Register müssen sie zusammen die Aufgabe lösen, zehn Unterschiede in einer Zeichnung zu finden, ohne dass der jeweils andere diese einsehen kann (Diapix-Aufgabe, vgl. Baker und Hazan, 2011). Entsprechend ist der Verlauf des freien Dialogs wahrscheinlich stärker von der spontanen Sprecherinteraktion geprägt als der der Diapix-Dialoge. Der freie Dialog ermöglicht den Sprecher\*innen aufgrund der offenen Instruktion, über Aktivitäten des vergangenen oder zukünftigen Wochenende zu sprechen, Erzähleinheiten zu produzieren, die sich in Form und Inhalt frei gestalten lassen.<sup>6</sup> Die Sprecher\*innen müssen den gesamten semantischen Rahmen planen und können hier recht frei den thematischen und auch funktionalen Fokus ändern. Weiterhin können die freien Dialoge eher dialogischer (verhältnismäßig gleiche Redeanteile und Sprecherwechsel) oder monologischer Natur (monologische Erzählweise einzelner Sprecher\*innen) sein, während die Diapix-Aufgaben beide Sprecher\*innen zur Interaktion zwingen (u. a. Bestätigung, Nachfragen). Bei häufigen Sprecherwechseln, Frage-Antwort-Interaktionen sowie Backchanneling sind möglicherweise eher spontansprachliche Mittel erforderlich als bei eher monologischen Erzähleinheiten einzelner Sprecher\*innen, was mit unterschiedlich distribuierten Flüssigkeitsmerkmalen je Abschnitt einherginge. In der Diapix-Aufgabe ist zusätzlich

---

<sup>6</sup> So können beispielweise verschiedene Erzählmodi (Tempus, Nacherzählung, direkte Rede etc.) gewählt werden.

neben dem Thema auch das lexikalische Repertoire ein Stück weit festgelegt, da diese zur Verbalisierung der Unterschiede benötigt wird (vgl. *Fleischer* in Beispiel 1). Die freie Auswahl aus dem bisher gelernten Vokabular ist für Lerner\*innen daher eingeschränkt und kann zu reduzierten Flüssigkeitsmaßen führen.

In dieser Studie untersuchen wir die Distribution und Aussagekraft von Flüssigkeitsmaßen im L2-Kontext zwischen zwei Registern und ergänzen dies mit einer Intraregisterperspektive, die die Untersuchung der Maße auf Abschnitte der einzelnen Äußerungskontexte bezieht. Wir kontrollieren darüber hinaus die Ergebnisse auch auf sprecherindividuelle Einflussfaktoren. Im nachfolgenden Abschnitt stellen wir unsere Datengrundlage sowie die untersuchten quantitativen und qualitativen Abschnitte vor und motivieren die Analysemethode.

## 3 Methode

### 3.1 Korpus

Wir verwenden das Wrocław Dialogue Corpus (WroDiaCo), ein gesprochenes, zeitaligniertes Lernerkorpus (Granger et al., 2015, Wisniewski, in diesem Heft), welches als WroDiaCo Version 2 die hier vorgenommenen Annotationen enthält (Wesolek et al., 2021). Entsprechend den in Hirschmann und Schmidt (in diesem Heft) diskutierten Desiderata für gesprochene Lernerkorpora steht es für die Forschung zur Verfügung<sup>7</sup> und ist ausführlich dokumentiert (vgl. Wesolek & Belz, 2021). Zudem ist ein Vergleich mit dem muttersprachlichen und methodisch gleich aufgebauten Korpus BeDiaCo (Belz et al., 2021) möglich. Das Korpus enthält acht Dialoge von 16 polnischen Deutschlerner\*innen mit einer DaF-Erfahrung von mindestens vier Jahren, wobei die Mehrheit der Versuchspersonen einen deutlich längeren Erwerb aufweist (vgl. Tabelle 1). Zusätzlich zur Erwerbsdauer gaben die Versuchspersonen eine Selbsteinschätzung ihres Sprachniveaus unter Berücksichtigung abgelegter Prüfungen im Rahmen des GER ab (Hilpisch, 2012).

In jeder der acht Dyaden (a, c, d, e, f, g, h, i) unterhalten sich zwei Versuchspersonen über das vergangene oder zukünftige Wochenende (themenorientiertes Register, Zeitbegrenzung vier Minuten). Der Rahmen wurde von der Experimentatorin im Vorfeld erklärt:

„Sie dürfen sich jetzt circa 4 min lang über ein Thema Ihrer Wahl, wie zum Beispiel über Ihr vergangenes oder Ihr kommendes Wochenende, unterhalten.“

<sup>7</sup> Das Korpus wird Zwecke über die Phonetikinstanz des Medienrepositoriums der Humboldt-Universität zu Berlin zur Verfügung gestellt (<https://rs.cms.hu-berlin.de/phon/>).

Anschließend lösen die Paare eine Diapixaufgabe (aufgabenorientiertes Register, keine Zeitbegrenzung). Hierfür erhalten beide Versuchspersonen eine unterschiedliche Version derselben Zeichnung mit dem Ziel, die eingebauten zehn Unterschiede zu finden, ohne dabei das Bild des Gesprächspartners sehen zu können (vgl. Baker & Hazan, 2011). Beide Register sind somit *a priori* durch die Art des Dialogs (themenorientiert vs. aufgabenorientiert) definiert.

**Tab. 1:** Übersicht über die Versuchspersonen mit Erwerbsdauer und geschätztem Niveau im Rahmen des GER in WroDiaCo v.2.

VP <sup>a</sup>	a1f	a2m	c1f	c2f	d1f	d2f	e1f	e2m
Jahre DaF	14	9	14	4	7,5	7	7	7
Niveau	C1	B1	k. A.	B2	C1	B1	B1	B1
VP <sup>a</sup>	f1f	f2f	g1m	g2f	h1f	h2f	i1f	i2m
Jahre DaF	18	9	11	13	10	15	7	13
Niveau	B2	B1	C1	C1	C1	C1	C1	B1

<sup>a</sup> Der letzte Buchstabe bezeichnet das Geschlecht (f = feminin, m = maskulin).

Die Dialogpaare kennen sich aus dem Studium und saßen sich gegenüber. Die Aufnahmen fanden in einem schallgedämpften Raum mit Nackenbügelmikrofonen und einem mobilen Aufnahmegerät statt. Weitere Informationen zum Ablauf des Experiments und zur Aufbereitung des Korpus finden sich in der Korpusdokumentation (Wesolek & Belz, 2021). Insgesamt enthält WroDiaCo 4,3 Stunden Sprachaufnahmen und 20 819 Tokens. Tabelle 2 gibt einen Überblick über Dauer und Tokenanzahl exklusive Pausen, unverständlichen Abschnitten und extralinguistischen Phänomenen wie Lachen oder Räuspern.

**Tab. 2:** Artikulationsdauer (min) und Tokenanzahl (auf der Annotationsebene *ortflue-1/-2* basierend) ohne Pausen, Unverständlichem und extralinguistischen Phänomenen je Dyade für Diapixe und freie Dialoge in WroDiaCo v.2.

	Dyade	a	c	d	e	f	g	h	i	Σ
Artikulations- dauer (min)	Diapix Frei	5,2	4,4	4,9	5,3	2,8	6,5	4,9	4,9	38,9
		1,7	1,7	1,4	1,5	0,8	1,9	1,7	1,9	12,6
Tokenanzahl	Diapix Frei	1 477	1 270	1 226	1 519	747	1 887	1 711	1 533	11 370
		496	460	328	427	188	600	554	533	3 586

### 3.2 Annotation

Unter Annotationen verstehen wir Interpretationen und explizite Zuweisungen von Kategorien zu einzelnen oder mehreren Einheiten in einem Korpus (Kuebler & Zinsmeister, 2015; Lüdeling, 2007). Annotationen sind also Repräsentationen strukturierten Wissens und stützen das theoretische Fundament unserer Analyse. WroDiaCo v.1 enthält eine Mehr-Ebenen-Architektur (u. a. Transliterationen), die wir in v.2 mit den für unsere Forschungsfrage relevanten Ebenen ergänzen. Hierzu korrigieren wir die Transliteration (Ebene *ortflue*) und fügen Pausen und Füllpartikeln (Ebene *flue*) hinzu sowie zwei Ebenen mit qualitativ definierten Abschnitten, auf denen Dialogzüge (Ebene *dia*) und gefundene Unterschiede in der Diapixaufgabe (Ebene *auab*) annotiert werden. Diese Abschnitte beruhen auf einer theoretischen Modellierung. Eine weitere Ebene enthält vier quantitativ definierte Abschnitte, die den Dialog in vier Teil teilt (Ebene *quab*) hinzu. Die Annotation erfolgt in Praat (Boersma & Weenink, 2019). Der folgende Auszug in Tabelle 3 zeigt die für diese Studie verwendeten Annotationen.<sup>8</sup>

**Tab. 3:** Annotationsbeispiel in WroDiaCo v.2 aus Dyade i, Diapix, Sek. 210–220, mit den Ebenen für die Transliteration *ortflue*, Pausen und Füllpartikeln *flue*, Dialogzügabschnitte *dia*, Aufgabenabschnitte *auab* und quantitative Abschnitte *quab*. Die Ziffern stehen für den Kanal der sprechenden Versuchsperson im Dialog.

<b>ort flue-1</b>		<b>ha</b>		<b>das gibt es nicht hier</b>		<b>dann vierter</b>		
flue-1	p		p		p		p	
dia-1	s	e					s	
auab-1-2	a4							
ort flue-2	äh	ein Text Fleischer			okay		also vielleicht haben wir vier	Unterschiede
flue-2	fp		p	<usb>	p	p		p
dia-2	aw		s		b			
quab-1-2	b							

<sup>8</sup> Das Beispiel ist qualitativ wiedergegeben, ohne Rücksichtnahme auf die Zeitalignierung.

Auf der Ebene *flue* sind Füllpartikeln (*fp*) und Pausen annotiert. Pausen erhalten den Wert *p* innerhalb und den Wert *s* zwischen den Abschnitten auf *dia*. Mit der Kategorie *s* umgehen wir die Entscheidung, diese Pause zum vorangegangenen Dialogabschnitt oder zum nachfolgenden Dialogabschnitt zuordnen zu müssen.

Die Annotationsebene *dia* enthält Spannenannotationen für Dialogzüge, die eine basale Unterteilung des sprachlichen Materials in Abschnitten zu Fragen (*fe* Entscheidungsfragen, *fw* offene Fragen), Antworten (*ap* positive Antworten, *an* negative Antworten, *au* unklare Antworten, *aw* offene Antworten), Erzähleinheiten (*e*) und Backchanneling (*b*) repräsentiert. Diese Einteilung basiert auf den Dialogstrukturannotationen von Carletta et al. (1997).

Die funktionale Abschnittsebene *auab* wird nur für die Diapixe und nur einmal je Dyade annotiert. Sie enthält die jeweiligen Zeitspannen, welche die Sprecher\*innen benötigten, um den nächsten Unterschied in ihren verschiedenen Bildern zu finden.

Die Ebene *quab* teilt die Dialoge in vier quantitativ definierte gleichlange Abschnitte in Abhängigkeit von der Gesamtlänge des Dialogs ein, um ein rein abstraktes Abschnittsmaß als Kontroll- und Vergleichsmaß zu definieren.

### 3.3 Auswertung

Die Audio- und TextGrid-Daten werden je Versuchsperson mit dem Paket *emuR* (Winlelmann et al., 2020) in das EMU-Datenbankformat konvertiert und können so in R (RCore Team, 2019) abgefragt und analysiert werden. Silben werden mit dem R-Paket *syly* v.0.1.5 berechnet (Michalke, 2017). und beruhen auf der diplomatischen Transliterationsebene (*ortflue*).

Personengebundenen sprachliche Flüssigkeitsmaße werden wie folgt ermittelt. Die Artikulationsgeschwindigkeit wird in Silben/Sekunde ( $\sigma/s$ ) je Sprecher\*in berechnet, ohne Pausen. Die Pausenanzahl enthält alle Werte, die mit *p* auf *flue* annotiert sind, Füllpartikeln werden anhand von *fp* gezählt.

Für die Beantwortung der ersten Forschungsfrage untersuchen wir den Effekt der unabhängigen Variablen Register und Dialogstruktur sowie ihrer Interaktionen auf die abhängigen Variablen Artikulationsgeschwindigkeit, Pausenanzahl und Füllpartikelanzahl *in toto* und *in parte* je Dialog mithilfe linearer gemischter Modelle (R-Pakete *lme4*, Bates et al., 2014 und *lmerTest*, Kuznetsova et al., 2017). Versuchspersonen werden als *random intercepts* modelliert.

Um die einzelnen Effekte der verschiedenen Faktorstufen auf die abhängigen Variablen zu ermitteln, werden *Post-hoc*-Vergleiche mit dem R-Paket *emmeans* berechnet. Das  $\alpha$ -Niveau ist hierbei Tukey-korrigiert für multiple Vergleiche (Lenth, 2019). Ferner werden marginale ( $R_m^2$ ) und konditionale ( $R_c^2$ ) Determina-

tionskoeffizienten als Gütekriterium der Regression angegeben (*MuMIn*, Barton, 2018).  $R_m^2$  erklärt dabei nur die Varianz für die unabhängigen Faktoren,  $R_c^2$  die Varianz der unabhängigen und zufälligen Faktoren gemeinsam (Nakagawa & Schielzeth, 2013).

Kapitel 4 gliedert sich nach den auszuwertenden Annotationsebenen *quab*, *dia* und *auab* auf und integriert jeweils den Vergleich der In-toto-Maße.

## 4 Ergebnisse

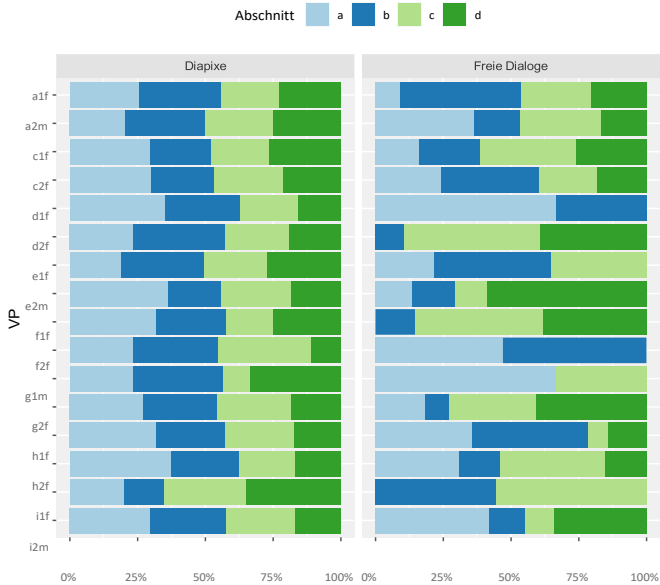
### 4.1 Quantitative Abschnitte

Vor der Auswertung der Flüssigkeitsmaße betrachten wir die Tokenanteile je Sprecher\*in für jeden der vier gleichlangen Dialogteile in beiden Registern. Während die relativen Tokenanteile in den Diapix-Dialogen ungefähr ausgeglichen sind, zeigen die freien Dialogen teils starke Schwankungen (vgl. Abbildung 1). Das deutlichste Ungleichgewicht ist für Dyade *f* zu beobachten, in der in der ersten Hälfte des Dialogs vornehmlich Sprecherin *f2f*, in der zweiten Hälfte dann vornehmlich Sprecherin *f1f* spricht.

Dies zeigt, dass die Register wie erwartet sprachlich verschieden funktionieren. Da die Diapix-Dialoge über die vier Abschnitte hinweg eine gleichmäßige Verteilung der Redeanteile zeigen, sind sie für die Anwendung von In-toto-Maßen (und auch von automatischen Samplinganalysen) besser geeignet als die freien Dialoge, die eine hohe Varianz aufweisen. Das heterogene Verhältnis der Tokenanteile je Versuchsperson in den themenorientierten freien Dialogen gibt Aufschluss darüber, dass die Interaktion nicht im gleichen Maße von Redewechseln geprägt ist wie in den Diapix-Dialogen. So scheinen bestimmte Dialoge (z. B. Dyade *f*) eher monologischen Charakter zu tragen.

Abbildung 2a zeigt, dass die Artikulationsgeschwindigkeit über alle Versuchspersonen *in toto* in den freien Dialogen geringfügig höher ist als in den Diapixen, mit einem Mittelwert von  $\bar{x} = 3,35$  Silben je Sekunde ( $\sigma/s$ ) und einer Standardabweichung von  $s = 0,56$  im Vergleich zu den Diapixen mit  $\bar{x} = 3,2$   $\sigma/s$ ,  $s = 0,51$ .<sup>9</sup>

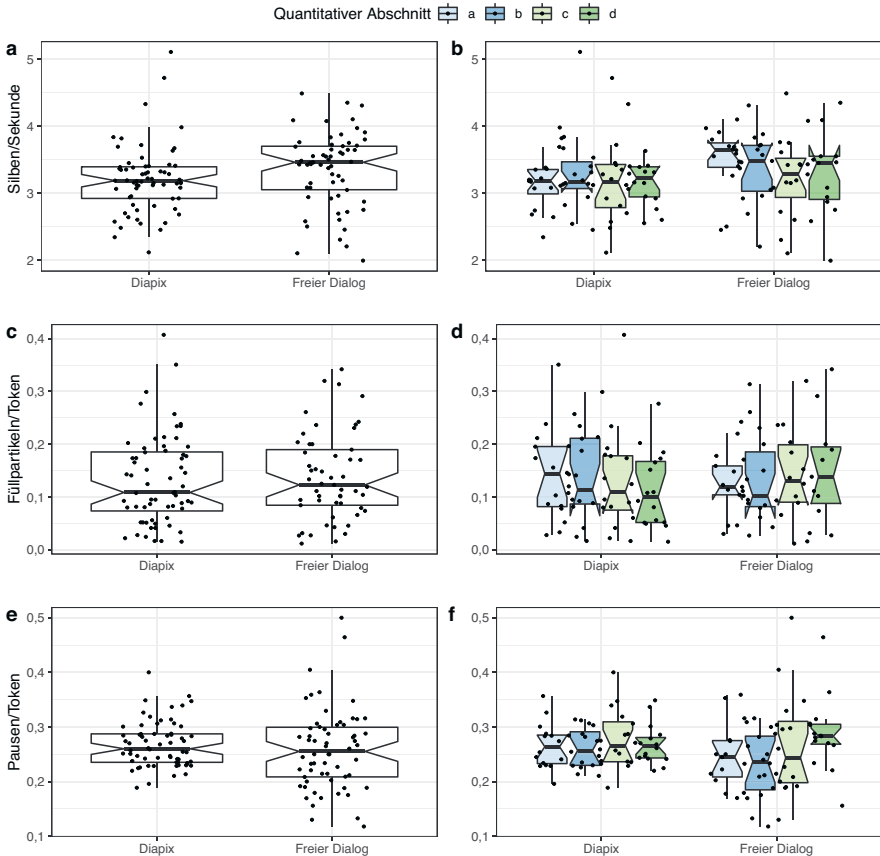
<sup>9</sup> Aus Abschnitt c (drittes Viertel) der Sprecherin *f2f* wurde ein Ausreißer von 9,6  $\sigma/s$  entfernt, der darauf zurückgeht, dass sie in diesem Abschnitt c nur das Wort *Familien* mit 4 Silben und einer Dauer von 0,42 Sekunden äußert, und so der anderen Sprecherin Hilfestellung bei der Suche nach der passenden Wortform gibt.



**Abb. 1:** Tokendistribution (basierend auf Annotationsebenen *ortflue-1/-2*) über die vier zeitlich gleichlangen Abschnitte a–d (basierend auf den Annotationsebenen *quab-1/-2*) für Versuchspersonen (VP) und Register in WroDiaCo v.2.

Ein lineares gemischtes Modell für die Artikulationsrate in Abhängigkeit von Register zeigt, dass der Unterschied in der Artikulationsrate zwischen den beiden Registern signifikant ist, wenn auch nicht besonders groß (vgl. Tabelle 4, erste Spalte). Aus Abbildung 2 b wird deutlich, dass dies besonders auf Dialogabschnitt a der freien Dialoge zurückgeht. Daraus leiten wir ab, dass bei jeder Analyse auf drei Ebenen verglichen werden muss: innerhalb einer sprachlichen Situation (Verlauf), zwischen verschiedenen sprachlichen Situationstypen (Registern) und zwischen den jeweils individuellen Situationen (jeder konkrete Dialog).

Für die relative Füllpartikelanzahl ergibt sich *in toto* kein signifikanter Unterschied zwischen den Registern (vgl. Abbildung 2c und Tabelle 5). Auch *in parte* zeigt sich keine Auffälligkeit zwischen den quantitativen Abschnitten – allenfalls ist ein sinkender Trend über den Verlauf der Diapixe zu beobachten (vgl. Abbildung 2d und Tabelle 5). Füllpartikeln sind in den freien Dialogen nicht gleichmäßig über die Redeanteile verteilt, während die sprecherspezifische Füllpartikeldistribution in den Diapixen über die Abschnitte hinweg relativ ähnlich bleibt (s. Abbildung 3). Im extremen Beispiel der Dyade f äußert Sprecherin *f2f* im dritten Dialogabschnitt nur ein Token und vierten Dialogabschnitt überhaupt



**Abb. 2:** Artikulationsgeschwindigkeit (a, b), Füllpartikelverhältnis (c, d) und Pausenverhältnis (e, f) *in toto* über alle Versuchspersonen je Register (linke Spalte) und *in parte* je quantitativem Abschnitt (basierend auf der Annotationsebene  $quab-1/-2$ ), rechte Spalte) in WroDiaCo v.2.

kein Token (s. Abbildung 1). Dementsprechend erwartbar beträgt auch die Anzahl ihrer Füllpartikeln (und damit ihr Flüssigkeitsmaß FP/Token) in diesen beiden Abschnitten null.

Die Pausenanzahl beträgt in den Diapixen 4 070 Pausen (26%), in den freien Dialogen 1 172 Pausen (23%). Am reinen Vorkommen sind die Register *in toto* nicht zu unterscheiden (vgl. auch Tabelle 4). Die über Versuchspersonen gemittelte Pausenanzahl je Token beträgt in den Diapixen im Mittel  $\bar{x} = 0,27$  ( $s = 0,04$ ), in den freien Dialogen  $\bar{x} = 0,26$  ( $s = 0,07$ ) (s. Abbildung 2e). Die Diapixe zeigen auch für dieses Maß keine Auffälligkeit über die vier Abschnitte hinweg. In den

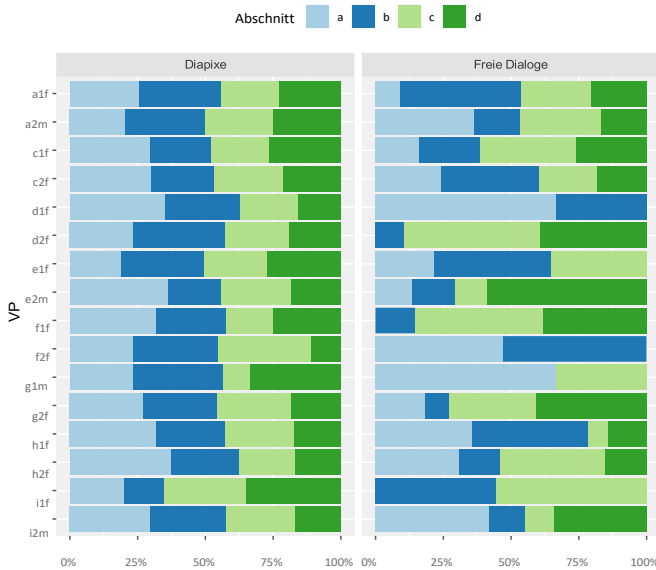


freien Dialogen zeigt Abschnitt d eine leicht erhöhte mittlere Pausenanzahl (s. Abbildung 2f).

**Tab. 4:** Lineare gemischte Modelle für die Artikulations-, Füllpartikel- und Pausenrate in Abhängigkeit von Register, mit Versuchspersonen als *random intercepts* und erklärter Varianz  $R^2$ .

	Art.-Rate	FP-Rate	Pausenrate
(Intercept)	3, 17 (0, 11)***	0, 25 (0, 02)***	0, 27 (0, 01)***
registerFrei	0, 12 (0, 05)*	0, 01 (0, 02)	-0, 01 (0, 01)
AIC	82, 80	203, 79	372, 64
Num. obs.	120	243	126
Num. groups: speaker	16	16	16
$R_m^2/R_c^2$	0, 02/0, 72	0/0, 14	0, 01/0, 46

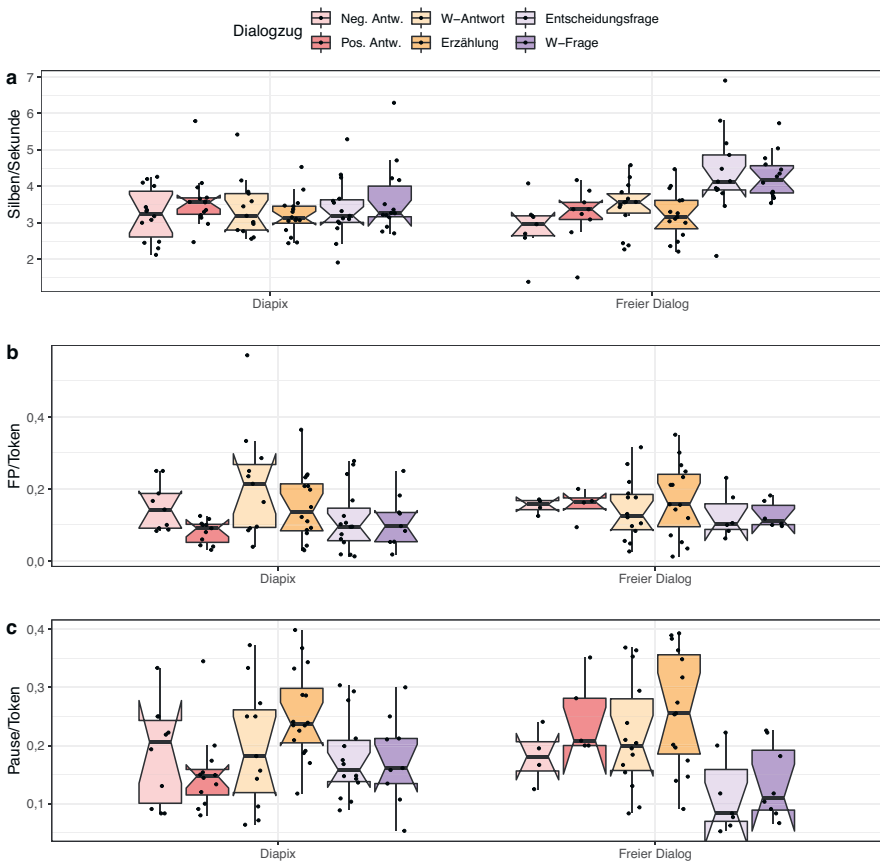
\*\*\*  $p < 0, 001$ , \*\*  $p < 0, 01$ , \*  $p < 0, 05$



**Abb. 3:** Füllpartikeldistribution über quantitative Abschnitte (basierend auf den Annotations-ebenen *quab-1/-2*) für Versuchspersonen (VP) und Register in WroDiaCo v.2.

**Tab. 5:** Mittelwert und Standardabweichung für Füllpartikel je Token über Sprecher\*innen *in parte* über die vier Abschnitte a–d in Diapixen und freien Dialogen.

Abschnitt	Diapix		Freie Dialoge	
	$\bar{x}$	s	$\bar{x}$	s
a	0,14	0,13	0,09	0,06
b	0,14	0,14	0,09	0,08
c	0,13	0,14	0,10	0,09
d	0,11	0,15	0,07	0,10



**Abb. 4:** Artikulationsgeschwindigkeit (a), Füllpartikelverhältnis (b) und Pausenverhältnis (c) *in parte* über alle Versuchspersonen je Register und Dialogzug (basierend auf den Annotations-ebenen dia-1/-2) in WroDiaCo v.2.

## 4.2 Dialogzüge

Im Folgenden wird keine rein quantitative, sondern eine an den qualitativ annotierten Abschnitten der Dialogzüge orientierte Segmentierung vorgenommen. Abbildung 4 a zeigt die Artikulationsgeschwindigkeit der Sprecher\*innen in beiden Registern über die Dialogzüge hinweg. Innerhalb des Diapixregisters zeigen sich keine signifikanten Unterschiede zwischen den Dialogzügen. Innerhalb der freien Dialoge wird in Entscheidungsfragen und W-Fragen schneller artikuliert als in negativen und positiven Antworten und in Erzählungen. In registerübergreifenden gleichen Kategorien zeigen nur Entscheidungsfragenabschnitte in Diapixen eine signifikant niedrigere Artikulationsgeschwindigkeit als in freien Dialogen (vgl. Abbildung 4 a und Tabelle 6).

Die Füllpartikelrate ist in W-Antworten der Diapixe signifikant höher als in positiven Antworten auf Entscheidungsfragen und signifikant höher als in Entscheidungsfragen (vgl. Abbildung 4 b und Tabelle 6ba–bb). Innerhalb der Dialogzugabschnitte der freien Dialoge und zwischen den Registern zeigt sich kein signifikanter Unterschied.

Die Pausenrate ist für Erzählungsabschnitte in den Diapixen signifikant höher als in positiven Antworten, Entscheidungsfragen und W-Fragen. Auch in den freien Dialogen ist die Pausenrate signifikant höher in den Erzählabschnitten als in Entscheidungs- und W-Fragen (vgl. Abbildung 4 c und Tabelle 6ca–cb). Registerübergreifend zeigt sich nur ein signifikanter Unterschied zwischen zwei verschiedenen Kategorien, nicht aber zwischen gleichen Kategorien.

**Tab. 6:** Signifikante Post-hoc-Effekte (Tukey-korrigiert für mehrfaches Testen mit 95 %-Konfidenzintervallen) und Intercepts der Modelle für a) Artikulationsrate ( $\sigma/s$ ), b) Füllpartikelrate (FP/Token) und c) Pausenrate in Abhängigkeit von Register (Diapix = Diapix-Dialoge, Frei = Themenorientierte Dialoge), Dialogzügen (Neg. Antw./Pos. Antw. = Negative und positive Antworten auf Entscheidungsfragen, W-Antw. = Antworten auf W-Fragen, E-Fr. = Entscheidungsfragen, Erz. = Erzählungen, W-Fr. = W-Fragen) und ihrer Interaktion mit Versuchspersonen als *random intercepts*.

Vergleich	$\beta$	s	df	Unt. KI	Ob. KI	t	p
<b>Artikulationsrate</b>							
Diapix,Neg. Antw. – Frei,E-Fr.	-1,08	0,28	133,16	-2,00	-0,16	-3,89	<0,01
Diapix,Neg. Antw. – Frei,W-Fr.	-0,99	0,27	132,88	-1,90	-0,09	-3,65	<0,05
Frei,Neg. Antw. – Frei,E-Fr.	-1,42	0,34	134,66	-2,55	-0,28	-4,16	<0,01
Frei,Neg. Antw. – Frei,W-Fr.	-1,33	0,34	134,15	-2,45	-0,21	-3,96	<0,01
Frei,Pos. Antw. – Frei,E-Fr.	-1,13	0,31	133,71	-2,18	-0,09	-3,61	<0,05
Frei,Pos. Antw. – Frei,W-Fr.	-1,05	0,31	133,27	-2,07	-0,02	-3,39	<0,05

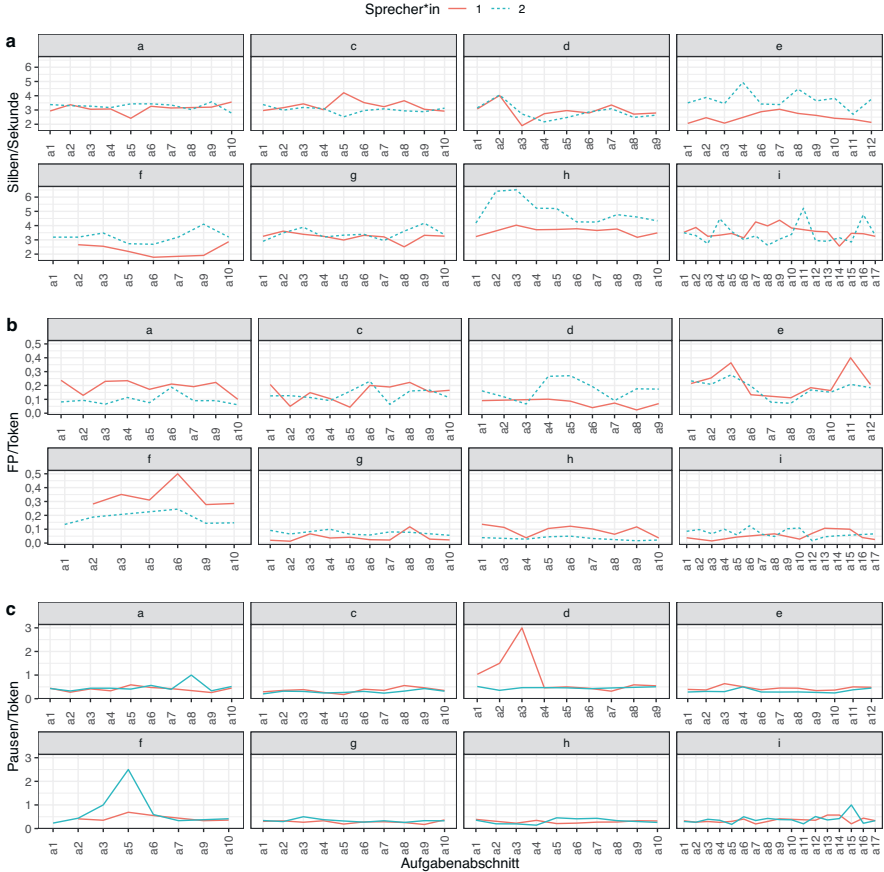
Tab. 6: (fortgesetzt)

Vergleich	$\beta$	s	df	Unt. KI	Ob. KI	t	p
<b>Artikulationsrate</b>							
Diapix,W-Antw. – Frei,E-Fr.	-0,96	0,28	132,06	-1,89	-0,02	-3,39	<0,05
Diapix,Erz. – Frei,E-Fr.	-1,13	0,27	133,62	-2,03	-0,23	-4,20	<0,01
Diapix,Erz. – Frei,W-Fr.	-1,04	0,26	133,40	-1,92	-0,17	-3,96	<0,01
Frei,Erz. – Frei,E-Fr.	-1,11	0,27	134,15	-2,02	-0,20	-4,06	<0,01
Frei,Erz. – Frei,W-Fr.	-1,02	0,27	133,91	-1,92	-0,13	-3,82	<0,05
Diapix,E-Fr. – Frei,E-Fr.	-0,96	0,27	132,73	-1,87	-0,05	-3,52	<0,05
<b>Füllpartikelrate</b>							
Diapix,Pos. Antw. – Diapix,W-Antw.	-0,14	0,03	94,74	-0,25	-0,04	-4,61	<0,001
Diapix,W-Antw. – Diapix,E-Fr.	0,10	0,03	94,34	0,01	0,20	3,58	<0,05
<b>Pausenrate</b>							
Diapix,Pos. Antw. – Diapix,Erz.	-0,09	0,03	100,35	-0,18	-0,00	-3,48	<0,05
Diapix,Pos. Antw. – Frei,Erz.	-0,10	0,03	100,80	-0,19	-0,01	-3,64	<0,05
Diapix,Erz. – Frei,E-Fr.	0,12	0,03	101,12	0,01	0,23	3,68	<0,05
Diapix,Erz. – Frei,W-Fr.	0,11	0,03	101,05	0,01	0,21	3,54	<0,05
Frei,Erz. – Frei,E-Fr.	0,12	0,03	101,09	0,02	0,23	3,83	<0,05
Frei,Erz. – Frei,W-Fr.	0,11	0,03	101,00	0,01	0,22	3,70	<0,05

### 4.3 Aufgabenabschnitte der Diapix-Aufgabe

Spezifisch für Diapix-Aufgaben ist das Finden von Bildunterschieden. Diese wurden als Abschnitte in der Ebene  $au_{ab}$  annotiert. Abbildung 5 a zeigt die Artikulationsgeschwindigkeit in den Diapixen über solche ergebnisorientierte, situational motivierten Abschnitte, die mit dem nächsten gefundenen Bildunterschied enden. Obwohl in den Diapixen genau zehn Unterschiede versteckt sind, finden die Versuchspersonenpaare zwischen neun und 17 Unterschiede. Da die Reihenfolge der Bildbesprechung nicht vorgegeben ist, sind die Abschnitte nicht zwischen den VP-Paaren vergleichbar. Die Analyse muss sich auf eine rein qualitative und abstrakte Beschreibung beschränken.

Augenfällige Unterschiede zwischen einzelnen Aufgabenabschnitten zeigen sich in der Variation der Artikulationsgeschwindigkeit je Sprecher\*in über verschiedene Abschnitte hinweg. Auch die Füllpartikelrate variiert teilweise stark (s. Abbildung 5 b). Möglicherweise liegen hier besondere Schwierigkeiten in der Interaktion oder der lexikalischen Beschreibung der Unterschiede vor. Für die Pausenrate lässt sich kaum eine Aussage treffen – die Sprecher\*innen produzieren fast immer ähnlich viele, was durch einige Ausreißer ergänzt wird (vgl. Abbildung 5 c).



**Abb. 5:** Artikulationsgeschwindigkeit (a), Füllpartikelfrequenz (b) und Pausenfrequenz (c) je Aufgabenabschnitt a1–a17 (basierend auf den Annotationsebenen auab-1/-2) und Sprecher\*in je Dyade a–i der Diapixe in WroDiaCo v.2.

## 5 Diskussion

Mit Blick auf unsere erste Forschungsfrage (Können abschnittsweise ermittelte Flüssigkeitsmaße andere oder komplementäre Aussagen über einen Dialog treffen als In-toto-Maße?) zeigen wir, dass dies zumindest für einige Flüssigkeitsmaße in Bezug auf die qualitativ ermittelten Dialogstrukturabschnitte gilt. Für die zweite Forschungsfrage (Können Flüssigkeitsmaße im Vergleich von *in parte* vs. *in toto* als Indikator herangezogen werden, um Register unterscheiden?) konnten

wir zeigen, dass mit In-toto-Maßen Register nur in einem von drei Maßen (Artikulationsgeschwindigkeit) und nur mit einer kleinen Effektstärke unterschieden werden können. Durch die Auswertung der In-parte-Maße hingegen konnte gezeigt werden, dass sich diese für die Artikulationsgeschwindigkeit in gleichen Kategorien nur für Entscheidungsfragen signifikant unterscheiden. Im Folgenden diskutieren wir die methodologischen Konsequenzen dieser Erkenntnisse.

Die erste Konsequenz ist, dass in Studien zu Flüssigkeitsmaßen getestet werden müsste, ob die gezählten Einheiten pro Zeiteinheit regelmäßig in Zeitabständen vorkommen können bzw. erwartbar sind. Gerade die freien Dialoge weisen sehr unterschiedliche Sprecher\*innenanteile auf und damit auch unterschiedliche Wahrscheinlichkeiten für Flüssigkeitsmerkmale im Verlauf des Dialogs. Zumindest eine Verlaufs- oder Abschnittsanalyse sollte heuristisch vor der Anwendung von In-toto-Maßen stehen, um die Adäquatheit der Analyse zu gewährleisten. Dies gilt insbesondere für Studien, die mit gesampelten Daten arbeiten, da einzelne Ausschnitte nicht zwingend etwas über den gesamten Dialog sagen können. Eine umfassendere Kenntnis des Datenmaterials scheint damit eine notwendige Bedingung für solche Studien zu sein.

Die zweite Konsequenz ist die Notwendigkeit einer theoretischen Einbettung der verwendeten Maße und der Einheiten, über die sie berechnet werden. In-toto-Maße können grundsätzlich ohne theoretische Einbettung berechnet werden. Damit sind sie jedoch nur bedingt aussagekräftig. Kategorien (Einheiten) setzen eine explizite oder implizite Modellierung und damit theoretische Annahmen voraus. In-parte-Maße hingegen müssen zwingend theoretisch eingebunden werden, da sie theoretische Konzepte wie beispielsweise aus der Dialogzugtheorie (Carletta et al., 1997), der Rhetorischen Strukturtheorie (Mann & Thompson, 1988) oder der Turn-Taking-Theorie (Sacks et al., 1974) in der Modellierung der Abschnitte umsetzen. Zwar gibt es schon Ansätze für In-parte-Maße in der Forschung zur sprachlichen Flüssigkeit, wie beispielsweise das Maß der mittleren Lauflänge, allerdings ist dieses oft arbiträr festgelegt, da es auf einer festzulegenden Mindestpausenlänge beruht. Ohne eine theoretische Einbindung dieser Mindestpausenlänge, zum Beispiel je nach Forschungsziel in Bezug auf produktive, artikulatorische Gegebenheiten oder Wissen über die Perzeption von Pausen kann auch das Maß selbst nicht interpretiert werden.

Für eine theoretisch modellierte Abschnittsdefinition spricht auch, dass man erst durch ein näheres Heranzoomen herausfinden kann, ob In-toto-Unterschiede ursächlich für eine bestimmte Kategorie sind. Diese Erkenntnis ist nicht neu – allerdings erlaubt es erst eine solche abschnittsweise Analyse, Auswertungen zu begründen und neue Vorhersagen aufzustellen. Am Beispiel der Entscheidungsfragen in den beiden untersuchten Registern lässt sich zeigen, dass wir nicht nur ein detailliertes Datenbild erhalten, sondern dass sich davon ausgehend leicht

neue Hypothesen für nachfolgende inhaltliche Forschungsfragen für die Lernerforschung aufstellen ließen. Für die Interpretation der quantitativen Abschnitte hingegen fehlt eine theoretische Basis.

Die In-parte-Auswertung der Aufgabenabschnitte hat jedoch gezeigt, dass auch eine theoretische Motivierung, die auf Registereigenschaften basiert, nicht immer zu aussagekräftigen Ergebnissen zwischen verschiedenen Abschnitten führen muss. Die Abschnitte waren für dieses Ziel möglicherweise zu grob gewählt und sind aufgrund der sehr unterschiedlichen Strategien der Sprecher\*innen und der gefundenen Unterschiede nicht vergleichbar. Hier müssten die gleichen Unterschiede in den gleichen Diapixen aufeinander abgebildet werden, um eine Aussage zu Besonderheiten mancher Aufgaben innerhalb dieses Registers treffen zu können.

Als Nebeneffekt der In-parte-Auswertung konnten wir feststellen, dass sich Diapixe besser zum Samplen eignen als freie Dialoge, da sie eher gleiche Redeanteile erwarten lassen. Mithilfe rein quantitativ modellierte Abschnitte lässt sich also die Samplinggeeignetheit eines Registers überprüfen, obwohl oder womöglich gerade da sie keine inhaltliche, sondern nur eine zeitliche Relation wiedergeben. Die Ergebnisse für die In-parte-Analyse mithilfe der quantitativen Abschnitte in den freien Dialoge lassen also vermuten, dass nicht alle Versuchspersonenpaare die interaktionale Erwartungshaltung des Registers erfüllen, da es zum Beispiel auch Dialoge mit eher monologischem Charakter enthalten kann. Diese Problematik muss von Registerstudien überprüft und berücksichtigt werden.

Wenn die explizite Modellierung der (Flüssigkeits-)Maßen zugrundeliegenden Abschnitte (beziehungsweise Einheiten) vorausgesetzt wird, ergeben sich unzählige Möglichkeiten, diesen Ansatz in verschiedene Richtungen auszubauen. So könnte die kategorielle Modellierung des Verlaufs durch die quantitativ ermittelten Abschnitte durch eine kontinuierliche Verlaufsanalyse erweitert werden, um feingranulare Intraregisterunterschiede festzustellen. Im Hinblick auf die qualitativen Abschnitte könnte die Mehr-Ebenen-Architektur stärker zu ihrem Vorteil verwendet werden, um zum Beispiel eine Kombination von Abschnitten auf verschiedenen Ebenen zu ermöglichen. Zukünftige Studien können sich diese abschnittswisen Vergleichsmöglichkeiten zunutze machen. Auch eine didaktische Einbettung der vorgestellten Analysen für bestimmte Abschnitte ist denkbar, obwohl es in diesem Bereich noch großen Aufbereitungsbedarf gibt (vgl. Tavakoli & Wright, 2020, S. 84–102).

Die bisher in der L2-Flüssigkeitsforschung verwendete In-toto-Analyse hat ihre Berechtigung für einen ersten Überblick und kann für den Vergleich strukturell ähnlich aufgebauter Datensätze sinnvoll sein. In-parte-Analyse ergänzen und erweitern diesen Ansatz. Damit ermöglicht die auf quantitativ und qualitativ modellierten Abschnitten beruhende Erschließung von Flüssigkeitsmaßen nicht

nur eine detailliertere Analyse von Lernerdaten, die zum Beispiel für die Registerforschung fruchtbar ist, sondern stellt auch insgesamt eine bessere Vergleichbarkeit heterogener Datensätze her.

## Literatur

- Baker, Rachel & Hazan, Valerie. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior research methods*, 43, 761–770. 10.3758/s13428-011-0075-y
- Baker-Smemoe, Wendy, Dewey, Dan P., Bown, Jennifer & Martinsen, Rob A. (2014). Does Measuring L2 Utterance Fluency Equal Measuring Overall L2 Proficiency? Evidence From Five Languages. *Foreign Language Annals*, 47, 707–728. 10.1111/flan.12110
- Barton, Kamil. (2018). MuMIn: Multi-Model Inference. Verfügbar 12. März 2019 unter <https://CRAN.R-project.org/package=MuMIn>
- Bates, Douglas, Maechler, Martin & Bolker, Ben (2014). lme4: Linear mixed-effects models using Eigen and S4. Verfügbar 28. September 2013 unter <http://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Belz, Malte. (2017). Glottal filled pauses in German. In R. Eklund & R. L. Rose (Hrsg.), *Proceedings of DiSS 2017: The 8th Workshop on Disfluency in Spontaneous Speech* (S. 5–8).
- Belz, Malte. (2021). *Die Phonetik von äh und ähm: Akustische Variation von Füllpartikeln im Deutschen*. J.B. Metzler. 10.1007/978-3-662-62812-6
- Belz, Malte, Mooshammer, Christine, Zöllner, Alina & Adam, Lea-Sophie. (2021). *Berlin Dialogue Corpus (BeDiaCo): Version 2*. <https://rs.cms.hu-berlin.de/phon>
- Belz, Malte, Sauer, Simon, Lüdeling, Anke & Mooshammer, Christine. (2017). Fluently disfluent? Pauses and repairs of advanced learners and native speakers of German. *International Journal of Learner Corpus Research*, 3, 118–148. 10.1075/ijlcr.3.2.02bel
- Belz, Malte & Trouvain, Jürgen. (2019). Are ‚silent‘ pauses always silent? 19. *International Congress of Phonetic Sciences (ICPhS)*. Verfügbar 21. August 2019 unter [https://icphs2019.org/icphs2019-fullpapers/pdf/full-paper\\_90.pdf](https://icphs2019.org/icphs2019-fullpapers/pdf/full-paper_90.pdf)
- Betz, Simon. (2020). Hesitations in Spoken Dialogue Systems. *Universität Bielefeld*. 10.4119/UNIBI/2942254
- Biber, Douglas. (2009). Multi-Dimensional Approaches. In A. Lüdeling & M. Kytö (Hrsg.), *Corpus Linguistics. An International Handbook* (S. 822–855). Mouton de Gruyter.
- Biber, Douglas & Conrad, Susan. (2009). *Register, genre, and style*. Univ. Press.
- Boersma, Paul & Weenink, David. (2019). Praat: doing phonetics by computer [Computer program]. Verfügbar 13. Juli 2019 unter <http://www.praat.org/>
- Bortfeld, Heather, Leon, Silvia D., Bloom, Jonathan E., Schober, Michael F. & Brennan, Susan E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, 44, 123–147. 10.1177/00238309010440020101
- Carletta, Jean, Isard, Amy, Isard, Stephen, Kowtko, Jacqueline, Doherty-Sneddon, Gwyneth & Anderson, Anne H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23, 13–31.
- Clark, Herbert H. & Fox Tree, Jean E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73–111. 10.1016/S0010-0277(02)00017-3



- Cucchiari, Catia, Strik, Helmer & Boves, Lou. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, 111, 2862–2873.
- Derwing, Tracey M. (2017). L2 Fluency Development. In S. Loewen & M. Sato (Hrsg.), *The Routledge handbook of instructed second language acquisition* (S. 246–259).
- Derwing, Tracy M., Munro, Murray J., Thomson, Ron I. & Rossiter, Marian J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533–557. 10.1017/S0272263109990015
- Derwing, Tracey M., Rossiter, Marian J., Munro, Murray J. & Thomson, Ron I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, 54, 655–679. 10.1111/j.1467-9922.2004.00282.x
- Fillmore, Charles. J. (1979). On fluency. In C. J. Fillmore, D. Kempler & W. S.-Y. Wang (Hrsg.), *Individual differences in language ability and language behavior* (S. 85–102). Academic Press.
- Foster, Pauline & Skehan, Peter. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 299–323.
- Gilquin, Gaëtanelle & Paquot, Magali. (2008). Too chatty. *English Text Construction*, 1, 41–61. 10.1075/etc.1.1.05gil
- Götz, Sandra. (2013). *Fluency in Native and Nonnative English Speech* (Bd. 53). John Benjamins Publishing Company. 10.1075/sc1.53
- Götz, S.andra (2019). Filled pauses across proficiency levels, L1s and learning context variables: A multivariate exploration of the *Trinity Lancaster Corpus Sample*. *International Journal of Learner Corpus Research*, 5, 159–180. 10.1075/ijlcr.17018.got
- Granger, Sylviane, Gilquin, Gaëtanelle & Meunier, Fanny. (2015). Introduction: learner corpus research – past, present and future. In S. Granger, G. Gilquin & F. Meunier (Hrsg.), *The Cambridge handbook of learner corpus research* (S. 1–5). Cambridge Univ. Press. 10.1017/CBO9781139649414.001
- Gut, Ulrike & Fuchs, Robert. (2017). Exploring speaker fluency with phonologically annotated ICE corpora. *World Englishes*, 36, 387–403. 10.1111/weng.12278
- Guz, Ewa. (2015). Establishing the fluency gap between native and non-native speech. *Research in Language*, 13, 230–247.
- Halliday, Michael A. K. (2014). *An introduction to functional grammar* (4. Aufl.). Routledge.
- Hilpisch, Kai. (2012). *Gemeinsamer Europäischer Referenzrahmen für Sprachen: Der GER im Überblick*. Diplomica.
- Hirschmann, Hagen & Schmidt, Thomas. (2022). Gesprochene Lernerkorpora: Methodisch-technische Aspekte der Erhebung, Erschließung und Nutzung. *Zeitschrift für germanistische Linguistik*.
- Kahng, Jimin. (2014). Exploring Utterance and Cognitive Fluency of L1 and L2 English Speakers: Temporal Measures and Stimulated Recall. *Language Learning*, 64, 809–854. 10.1111/lang.12084
- Karges, Katharina, Studer, Thomas & Hicks, Nina S. (2022). Lernersprache, Aufgabe und Modalität: Beobachtungen zu Texten aus dem Schweizer Lernerkorpus SWIKO. *Zeitschrift für germanistische Linguistik*.
- Kerz, Elma, Qiao, Yu, Wiechmann, Daniel & Ströbel, Marcus. (2020). Becoming Linguistically Mature: Modeling English and German Children's Writing Development Across School Grades. In J. Burstein, E. Kochmar, C. Leacock, N. Madnani, I. Pilán, H. Yannakoudakis & T. Zesch (Hrsg.), *Proceedings of the Fifteenth Workshop on Innovative Use of NLP for Building*

- Educational Applications* (S. 65–74). Association for Computational Linguistics. 10.18653/v1/2020.bea-1.6
- Klug, Katharina. (2013). „Ähm“ – Sind Häsitations-Partikeln sprecherspezifisch? Untersuchung der Parameter Grundfrequenz und Vokalqualität. In L. C. Anders, I. Bose, U. Hirschfeld & B. Neuber (Hrsg.), *Aktuelle Forschungsthemen der Sprechwissenschaft 3: Phonetik, Rhetorik, Sprechkunst, Sprach- und Stimmstörungen* (S. 65–94). Peter Lang.
- Kormos, Judit & Dénes, Mariann. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164. 10.1016/j.system.2004. 01.001
- Kuebler, Sandra & Zinsmeister, Heike. (2015). *Corpus Linguistics and Linguistically Annotated Corpora*. Bloomsbury Academic.
- Künzel, Hermann J. (1987). *Sprechererkennung: Grundzüge forensischer Sprachverarbeitung*. Kriminalistik Verlag.
- Kuznetsova, Alexandra, Brockhoff, Per B. & Christensen, Rune H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1–26. 10.18637/jss.v082.i13
- Labov, William. (1997). The Social Stratification of (r) in New York City Department Stores. *Sociolinguistics: A Reader* (S. 168–178). Nikolas Coupland; Adam Jaworski; Macmillan.
- Laver, John. (1994). *Principles of phonetics*. Cambridge University Press.
- Lay, Clarry H. & Paivio, Allan. (1969). The effects of task difficulty and anxiety on hesitations in speech. *Canadian journal of behavioural science*, 1, 25–37.
- Lennon, Paul. (1990). Investigating Fluency in EFL: A Quantitative Approach. *Language Learning*, 40, 387–417. 10.1111/j.1467-1770.1990.tb00669.x
- Lenth, Russel. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. Verfügbar 10. September 2019 unter <https://CRAN.R-project.org/package=emmeans>
- Lickley, Robin J. (2001). Dialogue moves and disfluency rates. *ITRW on Disfluency in Spontaneous Speech (DiSS'01)*, 93–96.
- Lickley, Robin J. (2015). Fluency and Disfluency. In M. A. Redford (Hrsg.), *The Handbook of Speech Production* (S. 445–469). John Wiley & Sons, Inc. 10.1002/9781118584156.ch20
- Lintunen, Pekka, Mutta, Maarit & Peltonen, Pauliina. (2020). Defining Fluency in L2 Learning and Use. In P. Lintunen, M. Mutta & P. Peltonen (Hrsg.), *Fluency in L2 Learning and Use* (S. 1–15). Multilingual Matters.
- Lüdeling, Anke. (2007). Das Zusammenspiel von qualitativen und quantitativen Methoden in der Korpuslinguistik. In G. Zifonun & W. Kallmeyer (Hrsg.), *Jahrbuch des Instituts für deutsche Sprache* (S. 28–48). De Gruyter.
- Mann, William C. & Thompson, Sandra A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8, 243–281. 10.1515/text.1.1988.8.3.243
- McDougall, K. & Duckworth, M. (2018). Individual patterns of disfluency across speaking styles: a forensic phonetic investigation of Standard Southern British English. *The International Journal of Speech, Language and the Law*, 25, 205–230.
- Michalke, Meik. (2017). sylly.de: Language Support for 'syllly' Package: German [(Version 0.1–2)]. Verfügbar 28. August 2019 unter <https://github.com/unDocUmeantIt/syllly>
- Moisl, Hermann. (2009). Exploratory Multivariate Analysis. In A. Lüdeling & M. Kytö (Hrsg.), *Corpus Linguistics* (S. 874–899). De Gruyter.
- Moniz, Helena, Batista, Fernando, Mata, Ana I. & Trancoso, Isabel. (2014). Speaking style effects in the production of disfluencies. *Speech Communication*, 65, 20–35. 10.1016/j.specom.2014.05.004

- Nakagawa, Shinichi & Schielzeth, Holger. (2013). A general and simple method for obtaining  $R^2$  from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. 10.1111/j.2041-210x.2012.00261.x
- Odebrecht, Carolin, Belz, Malte, Zeldes, Amir, Lüdeling, Anke & Krause, Thomas. (2017). RIDGES Herbology – Designing a Diachronic Multi-Layer Corpus. *Language Resources and Evaluation*, 51, 695–725. 10.1007/s10579-016-9374-3
- R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.
- Reagan, Andrew J., Mitchell, Lewis, Kiley, Dilan, Danforth, Christopher M. & Dodds, Peter S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5, 175. 10.1140/epjds/s13688-016-0093-1
- Sacks, Harvey, Schegloff, Emanuel A. & Jefferson, Gail. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735. 10.2307/412243
- Schnelle, Gohar. (2018). *Funktional bedingte Variation in der Evangelienharmonie Otrfrids von Weißenburg*. (Magisterarb.). Humboldt-Universität zu Berlin, Sprach- und literaturwissenschaftliche Fakultät. <http://dx.doi.org/10.18452/19250>
- Segalowitz, Norman. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, 54, 71. 10.1515/iral-2016-9991
- Skehan, Peter. (2003). Task-based instruction. *Language Teaching*, 36, 1–14. 10.1017/S026144480200188X
- Tavakoli, Parvaneh. (2010). Pausing patterns: differences between L2 learners and native speakers. *ELT Journal*, 65, 71–79. 10.1093/elt/ccq020
- Tavakoli, Parvaneh. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, 54, 133–150. 10.1515/iral-2016-9994
- Tavakoli, Parvaneh & Hunter, Ann-Marie. (2018). Is fluency being ‚neglected‘ in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research*, 22, 330–349. 10.1177/1362168817708462
- Tavakoli, Parvaneh. & Skehan, Peter. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Hrsg.), *Planning and Task Performance in a Second Language* (S. 239–273). John Benjamins Publishing Company.
- Tavakoli, Parvaneh. & Wright, Clare. (2020). *Second Language Speech Fluency*. Cambridge University Press. 10.1017/9781108589109
- Towell, Richard, Hawkins, Roger & Bazergui, Nives. (1996). The Development of Fluency in Advanced Learners of French. *Applied Linguistics*, 17, 84–119. 10.1093/applin/17.1.84
- Trofimovich, Pavel & Baker, Wendy. (2006). Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquisition*, 28. 10.1017/S0272263106060013
- Trouvain, Jürgen. (2022). Das IFCASL-Korpus als phonetisches Lernerkorpus. *Zeitschrift für germanistische Linguistik*.
- Trouvain, Jürgen & Truong, Khiet P. (2012). Comparing Non-Verbal Vocalisations in Conversational Speech Corpora. *4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals*, 36–39.
- van Hest, Erna, Poulisse, Nanda & Bongaerts, Theo. (1997). Self-Repair in L1 and L2 Production: an Overview. *ITL International Journal of Applied Linguistics*, 85–115.

- Wesolek, Sarah & Belz, Malte. (2021). Dokumentation und Annotationsrichtlinien für das Korpus WroDiaCo Version 2. *Humboldt-Universität zu Berlin*. 10.5281/zenodo.4541102
- Wesolek, Sarah, Belz, Malte. & Mooshammer, Christine. (2021). *Wrocław Dialogue Corpus (WroDiaCo): Version 2*. <https://rs.cms.hu-berlin.de/phon>
- Winkelmann, Raphael, Jaensch, Klaus, Cassidy, Steve & Harrington, Jonathan. (2020). *emuR: Main Package of the EMU Speech Database Management System* [R package version 2.1.1].
- Wisniewski, Katrin. (2022). Gesprochene Lernerkorpora des Deutschen: Eine Bestandsaufnahme. *Zeitschrift für germanistische Linguistik*.
- Włodarczak, Marcin & Wagner, Petra. (2013). Effects of talk-spurt silence boundary thresholds on distribution of gaps and overlaps. *Proceedings of Interspeech*, 1434–1437.