

Mittelgroße Lernerkorpora!

Feingranulare Registervariation – Parameter, Phänomene, Designentscheidungen

Anke Lüdeling

Milena Kuehnast / Beate Lütke / Nicole Schumacher / Samantha Weller

Humboldt-Universität zu Berlin

SFB 1412 *Register*, DFG-Referenznummer 416591334

DAKODA-Abschlusstagung, September 2025



*No man is an island,
Entire of itself.
Each is a piece of the continent,
A part of the main.
...
(John Donne, 1624)*

Das Folgende ist Ergebnis vieler, vieler Diskussionen und Kollaborationen mit Kolleg:innen.

Alphabetisch einige davon: Gaja Hartz, Hagen Hirschmann, Martin Klotz, Thomas Krause, Anna Shadrova

Wie kann der Erwerb von Registerwissen und Registerflexibilität korpusbasiert untersucht werden?

Wie kann der Erwerb von Registerwissen und Registerflexibilität korpusbasiert untersucht werden?

- Was ist Registerwissen?
- Was ist Registerflexibilität?
- Wie werden Registerwissen und Registerflexibilität erworben?
- **Welche Korpora braucht man?**
Welche Designentscheidungen muss man treffen?
Und: Wie sicher können wir uns sein, dass wir die richtigen Designentscheidungen treffen?

- 1 Registererwerb
- 2 Herausforderungen | Interindividuelle Variation
- 3 Herausforderungen | Parameter
- 4 Mitttelgroße Korpora

„Register means those aspects of socially recurring intraindividual variation in linguistic behavior that are influenced by situational and functional parameters.“
(Lüdelling et al., 2024)

Zu Register siehe unter vielen anderen Ágel und Hennig (2006), Biber und Conrad (2009), Biber und Gray (2010), Egbert und Gracheva (2022), Maas (2010), Matthiessen (2019) und Neumann (2014)

„Register means those aspects of socially recurring **intraindividual variation** in linguistic behavior that are influenced by situational and functional parameters.“
(Lüdelling et al., 2024)

Zu Register siehe unter vielen anderen Ágel und Hennig (2006), Biber und Conrad (2009), Biber und Gray (2010), Egbert und Gracheva (2022), Maas (2010), Matthiessen (2019) und Neumann (2014)

„Register means those aspects of **socially recurring** intraindividual variation in linguistic behavior that are influenced by situational and functional parameters.“
(Lüdeling et al., 2024)

Zu Register siehe unter vielen anderen Ágel und Hennig (2006), Biber und Conrad (2009), Biber und Gray (2010), Egbert und Gracheva (2022), Maas (2010), Matthiessen (2019) und Neumann (2014)

„Register means those aspects of socially recurring intraindividual variation in linguistic behavior **that are influenced by situational and functional parameters.**“
(Lüdelling et al., 2024)

Zu Register siehe unter vielen anderen Ágel und Hennig (2006), Biber und Conrad (2009), Biber und Gray (2010), Egbert und Gracheva (2022), Maas (2010), Matthiessen (2019) und Neumann (2014)

Register | Parameter und Phänomene

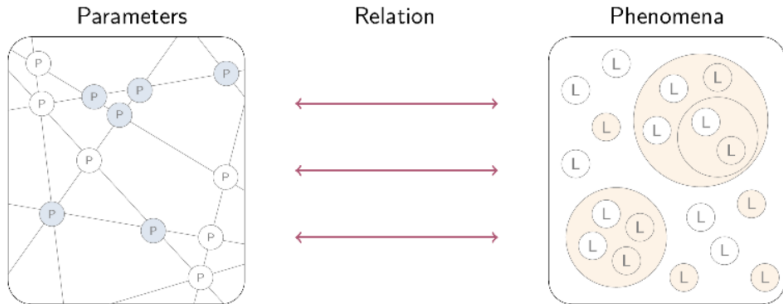


Abb. aus Lüdeling et al. (2024)

Parameter Die funktionalen und situationalen Parameter (sprachextern) sind bspw. Eigenschaften der Hörer:in, die Beziehung zwischen Sprecher:in und Hörer:in, Zweck des Diskurses, Lautstärke, Modus, ...

In vielen Registermodellen werden die Parameter als Liste dargestellt – wir sehen sie als ein komplexes Netz.

Parameter können einen jeweils unterschiedlichen Einfluss haben – eine wesentliche Frage ist dann, welcher Parameter welchen Einfluss hat.

Phänomene Sprachliche Phänomene auf allen Ebenen können von Registervariation betroffen sein. Seit Bibers Arbeiten (spätestens) werden ko-variiierende Mengen von Parametern untersucht.

Bisher passt das Modell gut zu Korpusuntersuchungen und es gibt auch Hunderte von Studien, die ein solches Modell annehmen, s.u.

Häufigste Vorgehensweise: Man erhebt/nutzt mehrere Korpora mit bestimmten (Konstellationen von) Parametern und untersucht und vergleicht dann die (ko-variiierenden) Phänomene.

Dann wird ein statistisches Modell (zur Dimensionsreduktion) angewendet und interpretiert.

Korpusstudien zu unterschiedlichen Sprachen, Phänomenen und mit unterschiedlichen statistischen Methoden, siehe bspw. Ädel (2008), Biber (2006a), Deshors (2015), Egbert und Gracheva (2022), Gilquin und Paquot (2008), Grieve et al. (2011) und Larsson (2019).

Bisher passt das Modell gut zu Korpusuntersuchungen und es gibt auch Hunderte von Studien, die ein solches Modell annehmen, s.u.

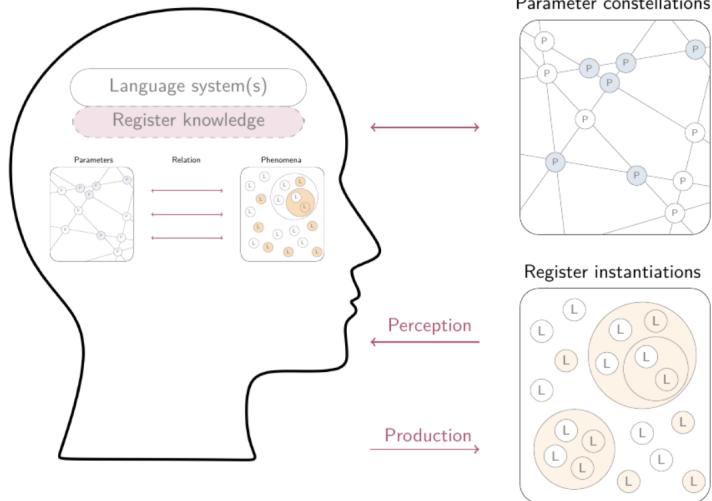
Häufigste Vorgehensweise: Man erhebt/nutzt mehrere Korpora mit bestimmten (Konstellationen von) Parametern und untersucht und vergleicht dann die (ko-variiierenden) Phänomene.

Dann wird ein statistisches Modell (zur Dimensionsreduktion) angewendet und interpretiert.

Aber es fehlt etwas Wichtiges.

Korpusstudien zu unterschiedlichen Sprachen, Phänomenen und mit unterschiedlichen statistischen Methoden, siehe bspw. Ädel (2008), Biber (2006a), Deshors (2015), Egbert und Gracheva (2022), Gilquin und Paquot (2008), Grieve et al. (2011) und Larsson (2019).

Was ist Registerwissen?



Wir gehen davon aus, dass der Erwerb von Registerwissen (und bestimmt auch der Verlust von Registerwissen) durch das ganze Leben hindurch stattfinden kann. Man muss jeweils eigene Kontexte/Hintergründe/kognitive Voraussetzungen etc. betrachten, es gibt sehr viele unterschiedliche Mehrsprachigkeitskonstellationen, hier drei Szenarien

- **im Kindesalter, L1:** intraindividuelle Variation muss 'entdeckt' werden, erste Variationsparameter müssen entschlüsselt werden, sprachliche Ausdrucksmöglichkeiten müssen erworben werden, mündlich
- **bei Adoleszenten/jungen Erwachsenen, L1:** argumentative und akademische Register, neu dazu: schriftlich, Erwerb von weiteren sprachlichen Ausdrucksmöglichkeiten, Erwerb von neuen funktionalen Anforderungen
- **in einer L2, gesteuerter Erwerb:** Transfer (auch vielleicht von Wissen über Parameter), Erwerb von sprachlichen Mitteln. Besondere Herausforderung: weniger sprachliche Erfahrung (weniger Situationen, generell weniger Möglichkeiten), Einflüsse von Lehrmaterialien

Was ist Registerflexibilität?

„the fact that children succeed with multiple registers shows that they understand the inherent flexibility of register; a speaker's social partners are subject to change and so are a speaker's registers.“ (Wagner et al., 2010, S. 1684)

„[...] Register Flexibility is defined as the ability to flexibly use a variety of linguistic resources – at the lexical, syntactic and discourse levels, with the awareness of which are the most appropriate for the communicative contexts at hand.“ (Qin & Uccelli, 2020)

„[...] register flexibility is understood as an individual productive skill which pertains to the capacity of the speaker for fast and controlled adjustments of language use based on sensitivity for changing communicative goals and circumstances [...]“ (Lüdelling et al., 2024, S. 24).

Wir brauchen also Korpora

- bei denen wir die Erhebungsparameter kennen und kodieren (als Metadaten (??))
- die idealerweise für dieselbe Sprecherin mehrere Register beinhalten (am besten in schnell wechselnden Situationen)

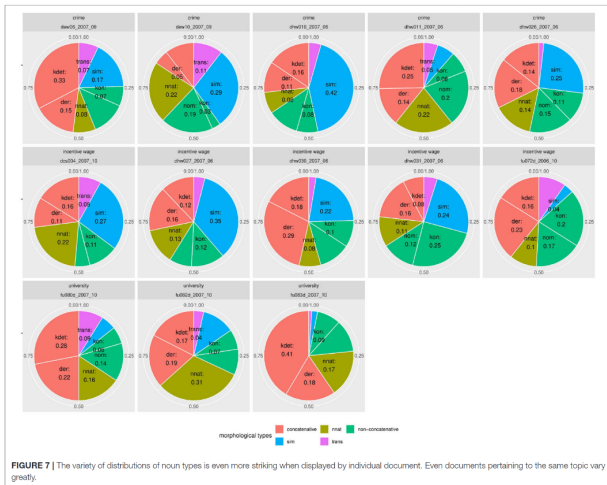
Aber wie können wir uns sicher sein, dass wir alles richtig interpretieren?

Wir (Projekt Co5 des SFB Register) bauen ein solches Korpus, das ReFlexAE-Korpus, siehe Kuehnast und Lütke (2022).

- 1 Registererwerb
- 2 Herausforderungen | Interindividuelle Variation
- 3 Herausforderungen | Parameter
- 4 Mitttelgroße Korpora

Eine der zugrundeliegenden Annahmen in quasi allen Registerstudien ist, dass sich Sprecher:innen in ähnlichen Situationen sprachlich ähnlich verhalten. Sprecher:innen lernen dann aus der Erfahrung.

Herausforderungen | Interindividuelle Variation



aus: Shadrova et al. (2021)

Anna Shadrova hat in verschiedene strukturiert erhobene Korpora auf lexikalische Überlappungen hin untersucht und findet viel mehr Variation als Überlappung (Shadrova, 2025, Abbildung S. 21).

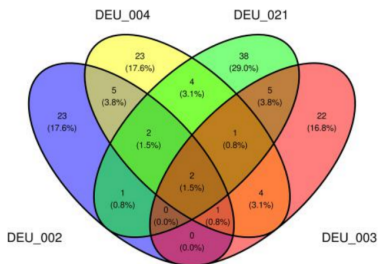


Figure 3: Venn diagram representing the overlap of verb lexemes between four texts in Kobalt (DEU_002, DEU_003, DEU_004, and DEU_021). Two verb lexemes occur in all four texts, 1, 1, 2, and 0 occur in different combinations of three texts, but not all four. 23, 23, 38, and 22 occur only in one of the texts, i.e., they are not shared with either of the other texts in the set. Percentages represent the proportion of the number of lexemes relative to the total number of lexemes in the set, not each speaker's individual set (the 23 lexemes that occur only in DEU_002 make up 17.6% of the total number of verb lexemes in the entire set of four documents). Sets of three texts are equivalent to the overlapping fields of three texts in this diagram.

Wir finden in ähnlichen Texten deutlich weniger Überlappung als erwartet – zumindest in lexikalischer Hinsicht (Lexeme, Wortbildungsmuster). Dies gilt nicht nur für Lerner:innen, sondern auch für LI-Sprecher:innen, also potentielle Inputgeber:innen.

„in light of the data presented, assumptions of self-similarity across levels of linguistic observation and abstraction cannot be considered valid for lexical distributions“ (Shadrova, 2025, S. 41)

Shadrova (2025) diskutiert theoretische Implikationen für Lexikontheorien und vor allem Pfadeffekte.

Hier nur die Frage: Könnte ein Teil der Variation auch daran liegen, dass die **Parameter** in den Erhebungssituationen unterschiedlich interpretiert werden? Dass also unterschiedliche Aufgaben erfüllt wurden?

- 1 Registererwerb
- 2 Herausforderungen | Interindividuelle Variation
- 3 Herausforderungen | Parameter**
- 4 Mitttelgroße Korpora

In Korpora werden Informationen zu Task, Erhebungssituation, Eigenschaften der Proband:innen in **Metadaten** kodiert. Dabei wurden im Laufe der Jahre immer mehr relevante Parameter entdeckt (task effects! social meaning!).

„For learner corpora to be maximally useful, it is essential that they come accompanied by more extensive metadata. Despite certain learner corpora being released with a decent set of learner and text variables, there remains a significant need for more comprehensive metadata and thorough corpus documentation.“
(Paquot, 2024, S. 572)

Aber welche Metadaten sind wichtig für Registervariation?

Registeruntersuchungen zeigen, dass vor allem Eigenschaften der Gesprächspartner:innen, die social role relationship und der Zweck der Interaktion eine Rolle spielen.

(Hier nicht im Fokus, aber: In welchen Metadatensteht viel zu den Eigenschaften der Gesprächspartner:innen?)

- Falko-Essays – narrative Teile in argumentativen Texten
- RUEG – unterschiedlicher Hang zum Spielen
- Transfer von Textstrukturen (Wan, 2023)
- kulturspezifische Unterschiede
- ...

„That-omission co-occurs with these stance features, suggesting that they are usually used in colloquial rather than **formal** registers.“ (Biber, 2006b, 237, unsere Hervorhebung)

„These studies also show that situational characteristics like level of preparation play a role in differences across registers of spoken discourse alongside **more traditionally recognized factors such as formality or informality**.“ (Goulart et al., 2020, 441, unsere Hervorhebung)

„The subdimension social distance encodes the interactants' relationship based on their mutual interactive history in terms of density as well as **formality of context**.“ (Neumann, 2014, 67, unsere Hervorhebung)

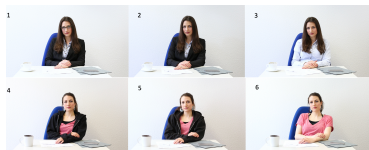
„That-omission co-occurs with these stance features, suggesting that they are usually used in colloquial rather than **formal** registers.“ (Biber, 2006b, 237, unsere Hervorhebung)

„These studies also show that situational characteristics like level of preparation play a role in differences across registers of spoken discourse alongside **more traditionally recognized factors such as formality or informality**.“ (Goulart et al., 2020, 441, unsere Hervorhebung)

„The subdimension social distance encodes the interactants' relationship based on their mutual interactive history in terms of density as well as **formality of context**.“ (Neumann, 2014, 67, unsere Hervorhebung)

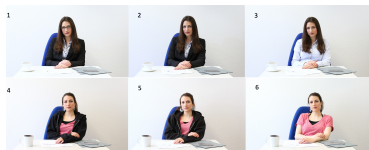
Haben Sie eine Intuition, was jeweils gemeint ist?

Wir denken, dass Formalität kein Primitiv ist, sondern ganz unterschiedliche Kombinationen von Eigenschaften (Parametern) beschreibt: Zweck, Ort, Eigenschaften der Gesprächspartner:innen, etc.



Abkürzung	Bedeutung
1ZBB	1 Zimmer mit B ad und B alkon
1ZKB	1 Zimmer mit K üche und B ad
1ZKBB	1 Zimmer mit K üche, B ad und B alkon
1ZKBT	1 Zimmer mit K üche, B ad und T errasse
1ZKDB	1 Zimmer mit K üche, D iele und B ad

Wir denken, dass Formalität kein Primitiv ist, sondern ganz unterschiedliche Kombinationen von Eigenschaften (Parametern) beschreibt: Zweck, Ort, Eigenschaften der Gesprächspartner:innen, etc.



Abkürzung	Bedeutung
1ZBB	1 Zimmer mit Bad und Balkon
1ZKB	1 Zimmer mit Küche und Bad
1ZKBB	1 Zimmer mit Küche, Bad und Balkon
1ZKBT	1 Zimmer mit Küche, Bad und Terrasse
1ZKDB	1 Zimmer mit Küche, Diele und Bad

Hier: Eine Ratingstudie zur Wahrnehmung und Bewusstheit von solchen Parametern (Kuehnast et al., n. d.). Wir fragen die VP also, was sie denken, das sie tun – das muss nicht notwendigerweise das sein, das sie tun.

Dazu gibt es im SFB Register noch weitere Untersuchungen (MeRAQ-Fragebögen - metalinguistic register awareness questionnaires).

SFB 1412

REGISTER



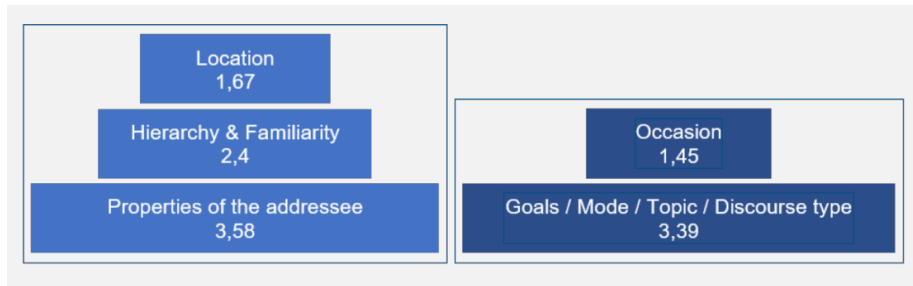
Task A: Situationelle Parameter für als Marker für **Formalität**

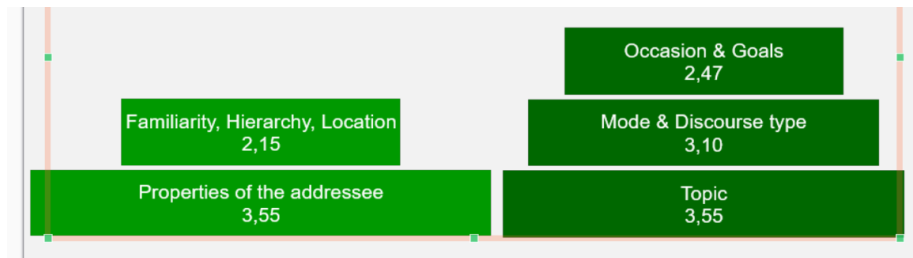
Instruktion: Um einzuschätzen, wie **formell eine Situation ist**, können verschiedene Situationsmerkmale eine Rolle spielen. Ordnen Sie diese Merkmale nach ihrer Wichtigkeit.

Task B: Situationelle Parameter als Marker für **Registeradaption**.

Instruktion: Verschiedene Situationsmerkmale beeinflussen **die Art und Weise, wie wir sprechen** oder schreiben. Ordnen Sie diese Merkmale nach ihrer Wichtigkeit.

- Beziehung zur anderen Person (soziale Hierarchie)
- Eigenschaften der anderen Person (z.B. Alter, Kleidung, Auftreten)
- Ort (z.B. Behörde, privates Umfeld)
- Kommunikative Ziele (Was möchte ich erreichen? Z.B. sich austauschen, Expertise & Fähigkeiten beweisen, flirten)
- Thema (Worum geht es?)
- Modus (i.e. gesprochen, geschrieben)
- Konkrete Kommunikationsform (z.B. Gespräch, E-Mail, Vortrag, Chatnachricht)
- Anlass der Kommunikation (Geburtstagsparty, Gerichtsverfahren, Vorstellungsgespräch, Mensabesuch)





Was bedeutet das?

Die Ratings der Proband:innen passen nicht zu den Ergebnissen der bisherigen Registerstudien. Das ist interessant für das Thema sprachliche Bewusstheit.

Was bedeutet das?

Die Ratings der Proband:innen passen nicht zu den Ergebnissen der bisherigen Registerstudien. Das ist interessant für das Thema sprachliche Bewusstheit.

Proband:innen finden die Kategorie **Anlass** am wichtigsten. Sie stellen sich wahrscheinlich einen 'typischen' solchen Anlass vor (also Geburtstag in der Familie und nicht der Geburtstag des Bundespräsidenten). Ein Anlass ist inhärent komplex - dazu gehören ein typischer Ort, typische Personen, ein typischer Zweck.

- Es ist vielleicht schwierig, nur eine Kategorie zu variieren; kann es sein, dass wir immer mit einem Ort/einer Person/einem Kleidungsstück einen Anlass verbinden und den bewerten?
- Die Vorstellung der Proband:innen ist stärker als das, was wir uns vorher denken (testen wir die Phantasie der VP?).

(Und: Formalität als Primitiv ist Mist.)

- 1 Registererwerb
- 2 Herausforderungen | Interindividuelle Variation
- 3 Herausforderungen | Parameter
- 4 Mittelgroße Korpora

Bestimmt gibt es Muster, die sich nur in sehr großen Datenmengen finden lassen (Verbstellungserwerb → DAKODA, task effects → Alexopoulou et al. (2017), syntaktische Komplexität → Vortrag von Meurers/Li etc.).

Aber es gibt auch Forschungsfragen, für die es unabdingbar erscheint, dass man sehr viel über die Korpora weiß.

Das bedeutet, dass die Korpora relativ klein sein müssen.

Andererseits wollen wir keine Einzelfälle betrachten.

Mittelweg?



**Der Mittel-Weg.
In Gefahr und grosser Noth
Bringt der Mittel-Weg den Tod.**



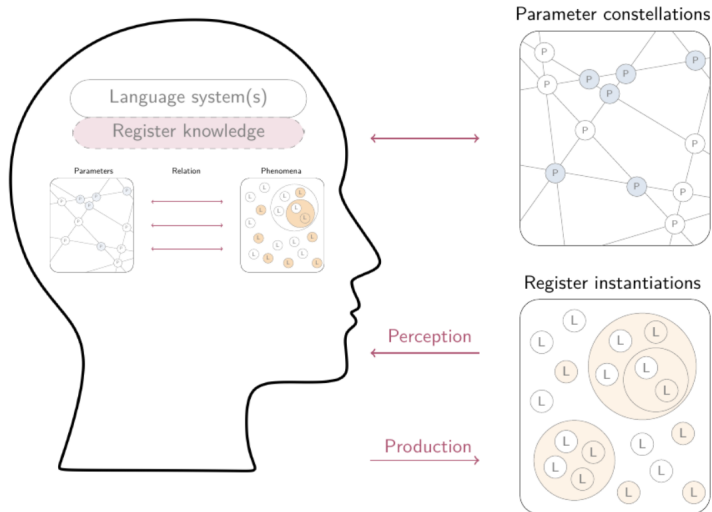
Μέτρον ἄριστον "der Mittelweg ist der Beste", zugeschrieben Kleobulos. "In Gefahr ..." aus Logau, Sämtliche Sinngedichte, 1872. Den Wein fand ich im Netz, getrunken habe ich ihn nicht. Postkarte vom Mittelweg in Leer/Ostfriesland nach einer Überschwemmung.

SFB 1412

REGISTER



Nochmal: Was ist Registerwissen?



Modell Wir brauchen ein **Registermodell**, in dem die Verarbeitung durch die Sprecherin mitmodelliert wird. Dann sind Korpusstudien vielleicht nicht immer adäquat. Mindestens müssen sie durch experimentelle Studien und andere Studien (wie die Ratingstudie) ergänzt werden.

Interindividuelle Variation Einige der (impliziten) Annahmen (Gleichartigkeit von Input in ähnlichen Parameterkonstellationen) scheinen so nicht zu gelten. Das ist insbesondere in der Lexik so. Kann es sein, dass Parameter unterschiedlich interpretiert werden?
Bei L2: Parameter können kulturspezifisch unterschiedlich interpretiert werden.

So ähnlich (allerdings nicht bezogen auf Registererwerb) argumentieren wir auch in Lüdeling et al. (2021) und bezogen auf Modellierung in Shadrova et al. (2025).

Zusammenfassung: Mittelgroße Korpora

Situationsspezifische Metadaten:

Übliches Vorgehen: Wir geben einen Kontext vor (task, Erhebungssituation) oder suchen Texte nach einem Kontext aus. Dabei verlassen wir uns darauf, dass die Proband:innen unsere Instruktionen und Parameter so interpretieren, wie wir uns das denken. Das muss aber nicht so sein.

→ Kann man dann Metadaten konsolidieren, um größere Korpora/
Korpussammlungen zu bauen?

Wir bräuchten mehr Forschung zum Verständnis der Metadaten – bei der Konsolidierung werden bereits abstrahierte Metadaten noch einmal abstrahiert. Was würde das bedeuten?

→ Wäre es nicht besser (Replikation mit unabhängigen Daten), mit mehreren mittelgroßen Korpora zu arbeiten, die Metadaten *nicht* zusammenzufassen, sondern die jeweiligen Besonderheiten ernst zu nehmen?

anke.luedeling@hu-berlin.de

<https://sfbi412.hu-berlin.de/>

<https://registergeknister.buzzsprout.com//>

- Ädel, A. (2008). Involvement features in writing: do time and interaction trump register awareness? In G. Gilquin, S. Papp & M. B. Diéz-Bedmar (Hrsg.), *Linking Up Contrastive and Learner Corpus Research* (S. 35–53). Rodopi.
- Ágel, V., & Hennig, M. (2006). Theorie des Nähe- und Distanzsprechens. In V. Ágel & M. Hennig (Hrsg.), *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650–2000*. (S. 3–31). Niemeyer.
- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: a large-scale learner corpus analysis employing Natural Language Processing techniques. *Language Learning*, 67(S1), 167–209.
- Biber, D. (2006a). Stance in Spoken and Written University Registers. *Journal of English for Academic Purposes*, 5(2), 97–116.
- Biber, D. (2006b). *University Language: A Corpus-Based Study of Spoken and Written Registers*. John Benjamins.
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge University Press.
- Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9(1), 2–20.
- Deshors, S. C. (2015). A Multifactorial Approach to Linguistic Structure in L2 Spoken and Written Registers. *Corpus Linguistics and Linguistic Theory*, 11(1), 19–50.



- Egbert, J., & Gracheva, M. (2022). Linguistic variation within registers: Granularity in textual units and situational parameters. *Corpus Linguistics and Linguistic Theory*, 19(1), 115–143. <https://doi.org/10.1515/cllt-2022-0034>
- Gilquin, G., & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, 1(1), 41–61–41–61.
- Goulart, L., Gray, B., Staples, S., Black, A., Shelton, A., Biber, D., Egbert, J., & Wizner, S. (2020). Linguistic Perspectives on Register. *Annual Review of Linguistics*, 435–455. <https://doi.org/10.1146/annurev-linguistics-011718-012644>
- Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2011). Variation Among Blogs: A Multi-dimensional Analysis. In *Genres on the Web* (S. 303–322). Springer.
- Kuehnast, M., & Lütke, B. (2022). Fachsprachliche Kompetenzentwicklung bei Grundschullehramtsstudierenden – metasprachliches Wissen und Registerflexibilität. In K. Madlener-Charpentier & G. Pagonis (Hrsg.), *Aufmerksamkeitslenkung und Bewusstmachung in der Sprachvermittlung. Kognitive und didaktische Perspektiven auf Deutsch als Erst-, Zweit- und Fremdsprache* (S. 371–393). Narr Francke Attempto Verlag. <https://doi.org/10.24053/9783772056871>
- Kuehnast, M., Weller, P. S., Schumacher, N., Lüdeling, A., & Lütke, B. (n. d.). Grasping Register Appropriateness – a Ranking Study of Metalinguistic Awareness.
- Larsson, T. (2019). Grammatical stance marking across registers. Revisiting the formal-informal dichotomy. *Register Studies*, 1(2), 243–268. <https://doi.org/https://doi.org/10.1075/rs.10004.lar>

- Lüdeling, A., Hirschmann, H., Shadrova, A., & Wan, S. (2021). Tiefe Analyse von Lernerkorpora. In H. Lobin, A. Witt & A. Woellstein (Hrsg.), *Deutsch in Europa. Sprachpolitisch - grammatisch - methodisch* (S. 235–283). de Gruyter.
- Lüdeling, A., Szucsich, L., Zeige, L. E., Adli, A., Alexiadou, A., Belz, M., Bouzouita, M., Bunk, O., Dreyer, M., Egg, M., Feulner, A. H., Fleischer, J., Gagarina, N., Hirsch, A., Jannedy, S., Knoeferle, P., Krause, T., Kutscher, S., Liu, M., ... Wiese, H. (2024). Register: Language Users' Knowledge of Situational-Functional Variation. Proposal frame text, CRC 1412 Register, Phase II. *REALIS: Register Aspects of Language in Situation*, 3. <https://doi.org/10.18452/29131>
- Maas, U. (2010). Literat und orat. Grundbegriffe der Analyse geschriebener und gesprochener Sprache. *Grazer Linguistische Studien*, 73, 21–150.
- Matthiessen, C. M. I. M. (2019). Register in Systemic Functional Linguistics. *Register Studies*, 1(1), 10–41. <https://doi.org/10.1075/rs.18010.mat>
- Neumann, S. (2014). *Contrastive Register Variation. A Quantitative Approach to the Comparison of English and German*. de Gruyter Mouton.
- Paquot, M. (2024). Learner Corpus Research: A Critical Appraisal and Roadmaps for Contributing (more) to SLA Research Agendas. *Corpus Linguistics and Linguistic Theory*, 20(3), 567–590. <https://doi.org/https://doi.org/10.1515/cllt-2024-0014>

- Qin, W., & Uccelli, P. (2020). Beyond Linguistic Complexity: Assessing Register Flexibility in EFL Writing Across Contexts. *Assessing Writing*, 45(100465).
<https://doi.org/https://doi.org/10.1016/j.asw.2020.100465>
- Shadrova, A. (2025). No Three Productions Alike: Lexical Variability, Situated Dynamics and Path Dependence in Task-based Corpora. *Open Linguistics*, 11, 1–58. <https://doi.org/10.1515/opli-2024-0036>
- Shadrova, A., Linscheid, P., Lukassek, J., Lüdeling, A., & Schneider, S. (2021). A Challenge for Contrastive L1/L2 Corpus Studies: Large Inter- and Intra-Individual Variation Across Morphological, but Not Global Syntactic Categories in Task-Based Corpus Data of a Homogeneous L1 German Group. *Frontiers in Psychology*, 12, 5267. <https://doi.org/10.3389/fpsyg.2021.716485>
- Shadrova, A., Lüdeling, A., Klotz, M., Hartz, G., & Krause, T. (2025). „Step away from the Computer!“: Über die linguistische Datenkategorisierung als Erkenntnisprozess und daraus folgende Herausforderungen bei der Nachnutzung von Annotationen und Annotationstools. *Zeitschrift für germanistische Linguistik*, (1), 166–214. <https://doi.org/https://doi.org/10.1515/zgl-2025-2005>
- Wagner, L., Greene-Havas, M., & Gillespie, R. (2010). Development in Children's Comprehension of Linguistic Register. *Child Development*, 81(6), 1678–1686.
- Wan, S. (2023). *Argumentationsstrategien chinesischer Deutschlerner/-innen*. J.B. Metzler.