



## Tagung

# „META-LING 2023 - Methodological Exploration and Technological Advances in Linguistics“

10.11.2023 an der Universität Bamberg

Carolin Cholotta & Christine Renker, Otto-Friedrich-Universität Bamberg  
BAGL - Bamberger Graduiertenschule für Linguistik.

## Tagungsprogramm

9:00-9:30	MG2/00.10 <i>Begrüßung</i>		
9:30-10:30	MG2/00.10 Keynote <i>Open Science in der Linguistik: Tipps, Tricks und Best Practices</i> Jun.-Prof. Dr. Stefan Hartmann (Universität Düsseldorf)		
10:30-10:45	MG2/00.10 <i>Kaffeepause</i>		
10:45-11:45	MG2/00.10 Workshop <i>Open science: Data publication using TROL-Ling</i> Dr. Lukas Sönning (Universität Bamberg)	MG1/00.09 Workshop <i>Innovative Datenerhebung mit PsychoPy</i> Mina Mikuljanac (Universität Bamberg) <b>entfällt!</b>	MG2/02.05 Workshop <i>Saubere, nachvollziehbare und sichere Datenerhaltung – nur wie?</i> <i>Tidy, comprehensible and secure data – but how?</i> Katharina Karges (Universität Bamberg)
11:45-13:00	<i>Mittagspause</i>		

<b>13:00-13:30</b>	<p>MG1/02.06 Vortrag</p> <p><i>Do it yourself Semantische Annotation – zur Auswertung mit dem Korpusanalysetool #LancsBox</i></p> <p>Miriam Schwebler (Universität Würzburg)</p>	<p>MG1/02.09 Vortrag</p> <p><i>Individual what? Measuring individual differences in psycholinguistic research</i></p> <p>Katja Haeuser (Universität des Saarlandes)</p>
<b>13:30-14:00</b>	<p>MG1/02.06 Vortrag</p> <p><i>Zur Distribution und Morphologie der althochdeutschen gi-Verben</i></p> <p>Marco Bruckmeier (Universität Bamberg)</p>	<p>MG1/02.09 Vortrag</p> <p><i>Enhancing Consistency and Reproducibility in Interval Measurements of Monophthong Vowels: A Novel Approach Using Glottal Pulse Units</i></p> <p>Nicholas Peterson (Universität Bamberg)</p>
<b>14:00-14:15</b>	<p>MG2/00.10 Kaffeepause</p>	
<b>14.15-14.45</b>	<p>MG1/02.06 Vortrag</p> <p><i>Deutsch-Chinesisches Lernerkorpus – DeChiLKo</i></p> <p>Ming Liu (Universität Göttingen)</p>	<p>MG1/02.09 Vortrag</p> <p><i>A Fundamental Shift in NLP? Traditional Models vs. Neural Networks and Large Language Models – Observations in Student Bachelor's and Master's Theses</i></p> <p>Robin Jegan (Universität Bamberg)</p>
<b>14:45-15:15</b>		<p>MG1/02.09 Vortrag</p> <p><i>From Byte to Babel: Large Language Models and the Tower of Linguistic Knowledge</i></p> <p>Bastian Bunzeck (Universität Bielefeld)</p>
<b>15:15-15:30</b>	<p>MG2/00.10 Kaffeepause</p>	
<b>15:30-16:00</b>	<p>MG2/00.10 Abschlussdiskussion</p> <p>Carolin Cholotta &amp; Christine Renker (Universität Bamberg)</p>	

## **Workshop: Innovative Datenerhebung mit PsychoPy**

Mina Mikuljanac

DiKuLe: Digitale Kulturen der Lehre entwickeln

Universität Bamberg

[mina.mikuljanac@uni-bamberg.de](mailto:mina.mikuljanac@uni-bamberg.de)

Der Workshop führt in die open-source Software PsychoPy ein, die eine Reihe von neuen Möglichkeiten der linguistischen Datenerhebung eröffnet: Ob Reaktionszeiten nach visuellem oder auditivem Input messen oder die Zuordnung von gesprochenen Stimuli zu dialektalen Karten mit Mouse- oder Eye-Tracker verfolgen: PsychoPy macht hybride Studiendesigns möglich, die zuvor mit großem Aufwand verbunden waren. Das Besondere dabei ist die intuitiv bedienbare Builder-Benutzeroberfläche, in der sich vieles ganz ohne Programmierkenntnisse umsetzen lässt. Und noch ein Plus: Mit PsychoPy lassen sich Studien auch online durchführen!

## Workshop: Open science: Data publication using TROLLing

Dr. Lukas Sönning

Lehrstuhl für Englische Sprachwissenschaft

Universität Bamberg

[lukas.soenning@uni-bamberg.de](mailto:lukas.soenning@uni-bamberg.de)

In this workshop, you will learn how to publish data on the Tromsø Repository of Language and Linguistics (TROLLing), a language data repository (<https://dataverse.no/dataverse/trolling>). We discuss the structure and components of a typical TROLLing post and look at how to write a ReadMe file, which describes the dataset in detail and therefore constitutes the heart of any TROLLing post. Examples for various types of language data (corpus/questionnaire/experimental data) will be provided. We will also walk through the submission and review process, and appreciate the linguistically-informed feedback and data curation services provided by the friendly TROLLing staff.

## **Workshop: Saubere, nachvollziehbare und sichere Datenhaltung – nur wie?**

### **Tidy, comprehensible and secure data – but how?**

Katharina Karges

Professur Deutsche Sprachwissenschaft / Deutsch als Fremdsprache

Universität Bamberg

[katharina.karges@uni-bamberg.de](mailto:katharina.karges@uni-bamberg.de)

Im Workshop wird ein Einblick in die Prinzipien von Forschungsdatenmanagement, Open Data und Versionsverwaltung gegeben. Wir werden uns mit Git und R Markdown beschäftigen und auf Wunsch auch den Package Manager renv besprechen. Kenntnisse in R sind für diesen Workshop hilfreich, aber keine Voraussetzung. Ein eigener Computer sollte mitgebracht werden. Wer alles live ausprobieren möchte, sollte bereits R, RStudio und GitHub installiert haben. Entsprechende Anleitungen sind im Internet zu finden.

The workshop will give an introduction into the principles of research data management, open data and version control. We will deal with Git and R Markdown and, if desired, also discuss the package manager renv. Knowledge of R is helpful but not a requirement. Please bring your own computer. If you want to try everything live, you should already have R, RStudio and GitHub installed. Helpful instructions can be found on the internet.

Workshop-Sprache ist deutsch (english on request).

## ***Do it yourself* Semantische Annotation**

### **zur Auswertung mit dem Korpusanalysetool #LancsBox**

Miriam Schwebler

Professur für deutsche Sprachwissenschaft  
Universität Würzburg  
[miriam.schwebler@uni-wuerzburg.de](mailto:miriam.schwebler@uni-wuerzburg.de)

Das Korpusanalysetool #LancsBox ermöglicht die Berechnung von Frequenzlisten, N-Grams und Kollokationen in großen, selbst erstellten (medial schriftlichen) Korpora. Die eingespeisten Texte werden automatisch lemmatisiert und mit POS-Tags versehen. An diesen Annotationsebenen können dann induktive Musterberechnungen und deduktive Korpusrecherchen ansetzen. Für verschiedene sprachwissenschaftliche Fragestellungen kann es jedoch notwendig sein, diese Annotationen durch semantische Informationen zu ergänzen.

Im Rahmen meines Promotionsprojektes habe ich die diskursive Konstruktion von Wissen über (geographischen) Raum untersucht. Ausgehend von der Hypothese, dass ein analytischer Zugriff auf diskursive Wissenskonstruktion über sprachliche Muster möglich ist, wurden induktiv sprachliche Muster in einem thematischen Korpus berechnet. Dabei waren natürlich vor allem solche Realisierungen der Muster interessant, die in einem direkten thematischen Zusammenhang mit (geographischen) Räumen stehen. #LancsBox ist jedoch nicht in der Lage, solche Muster von anderen Mustern zu unterscheiden, da das Analyseprogramm keinen Zugriff auf die Bedeutung der Muster hat.

Hier kann eine fragestellungsspezifische semantische Annotation ansetzen. Annotiert wurden in diesem Fall Substantive, die im jeweiligen Kontext auf Raum referieren. Erscheinen sie in einem Muster, handelt es sich mit großer Wahrscheinlichkeit um ein Muster, das für die Fragestellung interessant ist. Die sog. Placebezeichnungen wurden an einem Korpus-Subset qualitativ erhoben, wobei das Textanalysetool MAXQDA zum Einsatz kam. Anschließend wurde die Annotation über verschiedene Abstraktionsverfahren auf das Gesamtkorpus ausgeweitet. Das Ergebnis war ein annotiertes Korpus, in dem Lemma-, POS- und semantische Informationen kombiniert abgefragt werden können.

Der Vortrag will

- die Analyseprogramme #LancsBox und MAXQDA und ihre Verzahnung vorstellen,
- den Annotationsworkflow von der Identifikation der Placebezeichnungen bis zu ihrem Tagging im Korpus nachzeichnen und
- die Potenziale semantischer Annotation für Korpusanalysen herausarbeiten.

#### **Software:**

Březina, Václav/Pierre. Weill-Tessier/Anthony McEnery (2020), #LancsBox v. 5.x, Lancaster, UK: Lancaster University, verfügbar unter: <http://corpora.lancs.ac.uk/lancsbox/index.php>.  
VERBI Software (1989-2021), MAXQDA. Software für qualitative Datenanalyse, Berlin: Consult. Sozialforschung GmbH, verfügbar unter: <https://www.maxqda.com/de/>.

# Individual what? Measuring individual differences in psycholinguistic research

Katja Haeuser

SFB Information Density & Linguistic Encoding  
Universität des Saarlandes  
[khaeuser@coli.uni-saarland.de](mailto:khaeuser@coli.uni-saarland.de)

The past ten years or so have seen a dramatic increase in psycholinguistic research that investigates to what extent performance in linguistic tasks (e.g., online reading) is modulated by individual differences in cognitive functions, for example in working memory or inhibitory capacity. For example, processing of object-relative sentences has been reported to be more difficult for individuals with lower working memory capacity, presumably because the resource pool to maintain complex information in short-term memory is inherently more limited in these individuals.

However, the extant literature varies greatly regarding how precisely individual differences can be measured. That, together with the fact that spurious interactions between linguistic performance and individual differences emerge easily in studies with small sample sizes, has led to an explosion of individual difference tasks used in psycholinguistic research, with little guidance or insight for early-career researchers regarding which task to use best.

In this contribution, I will argue that individual differences are best measured at the same construct level the main linguistic task addresses. For example, if the main linguistic task measures processing of unpredictable nouns (a semantic task), a reliable individual difference task should also measure effects at the semantic level. I will provide examples for such tasks based on the practices in my lab, and I will show examples on how to preprocess and extract the relevant information from these tasks. My goal is to provide early-career researchers with useful tools to investigate individual differences and inspire critical thinking about the nature of individual difference tasks.

## Zur Distribution und Morphologie der althochdeutschen gi-Verben

Marco Bruckmeier

Lehrstuhl Deutsche Sprachwissenschaft  
Universität Bamberg  
[marco.bruckmeier@uni-bamberg.de](mailto:marco.bruckmeier@uni-bamberg.de)

Die althochdeutschen gi-Verben stellten in der historischen germanistischen Linguistik der vergangenen 150 Jahre eines der am intensivsten diskutierten Phänomene dar und noch immer widmen sich zahlreiche Veröffentlichungen dem Gegenstand. Neben der kategorialen Frage – Aspekt oder Aktionsart? – steht insbesondere zur Diskussion, wie systematisch und obligatorisch das gi-Präfix als Perfektivierer verwendet wurde. Eroms (1997) erkannte, dass ein Großteil der althochdeutschen Verben in einer unpräfigierten und einer gi-präfigierten Variante vorkommt. Diese Paarigkeit zweier Verben ist typisch für Aspektsprachen, in denen sich imperfektive und perfektive Verben gegenüberstehen und die Kodierung des Verbalaspektes übernehmen. Doch Eroms' Erkenntnis basiert lediglich auf Impressionen – aus diesem Grund widmet sich dieser Beitrag der Domäne der althochdeutschen gi-Verben mit einem umfassenden korpuslinguistischen Design. Mittels des Althochdeutschen Wörterbuchs (AWB Leipzig) wird ein Korpus der gi-Verben erstellt und das verbale Inventar erstmals sichtbar gemacht. Anschließend wird dieses hinsichtlich der Type-Distribution, der Verbklassenzugehörigkeit und der Paarigkeit zum Simplexverb beleuchtet. Distribution und morphologische Zusammensetzung liefern hierbei Aufschlüsse über den Stellenwert, den Status und das System der gi-Verben.

### **Literatur:**

Eroms, Hans-Werner (1997): Verbale Paarigkeit im Althochdeutschen und das 'Tempussystem' im 'Isidor'. Zeitschrift für das deutsche Altertum und deutsche Literatur 126(1), 1–31.



# Enhancing Consistency and Reproducibility in Interval Measurements of Monophthong Vowels: A Novel Approach Using Glottal Pulse Units

Nicholas Peterson

Kompetenzzentrum für Digitales Lehren und Lernen

Universität Bamberg

[nicholas.peterson@uni-bamberg.de](mailto:nicholas.peterson@uni-bamberg.de)

Manual interval measurements have long been a prevalent method for analyzing monophthong vowels (Di Paolo and Yaeger-Dror 2011). This involves measuring an interval in which the first, second, and sometimes more formants correspond to so-called “steady-state patterns”. Furthermore, this steady state is generally presumed to align with a target for monophthong vowels (Thomas 2011), and much research relies on the explanatory power of steady states and vowel targets as accurate measures of vowel identity (Rogers et al. 2013; Ladefoged 2003).

However, the above method suffers from the following drawbacks:

1. Potential bias due to wide researcher freedom in creating measurements
2. Future irreproducibility of processed data owing to inconsistencies in creating measurements

This contribution, firstly, investigates these drawbacks by presenting an original empirical annotation study involving four researchers with ample experience annotating phonetic data according to the above mentioned manual interval method. Using *Praat* (Boersma and Weenink 2023), the participants were asked to place 36 vowel interval measurements (9 vowel categories in total), unaware that each vowel token was duplicated three times to assess intra- and inter-participant consistency. Results suggest considerable inter- and intra-annotator variance for vowels with inherent dynamic qualities, while measurements for shorter static vowels show more consistency, which matches prevailing expectations (Kendall et al. 2021; Ladefoged 1996).

Secondly, this contribution offers a solution to these findings by introducing a novel method using *Python* (van Rossum and Drake 2023), which anchors measurement boundaries at glottal pulse intervals in *Praat*. Bias is thereby reduced by limiting annotator freedom, which assists in creating reproducible and consistent measurement intervals. A preliminary test using this approach has shown promising indications. By combining this novel method with general guidelines, this research proposes a potential solution to enhance the reproducibility

and comparability of processed phonetic data.

**References:**

- Boersma, Paul and David Weenink. 2023. *Praat: doing Phonetics by computer*.
- Di Paolo, Marianna and Malcah Yaeger-Dror (eds.) (2011). *Sociophonetics: A student's guide* (English Language and Linguistics). London, New York: Routledge Taylor & Francis Group.
- Kendall, Tyler, Charlotte Vaughn, Charlie Farrington, Kaylynn Gunter, Jaidan McLean, Chloe Tacata and Shelby Arnson. 2021. Considering Performance in the Automated and Manual Coding of Sociolinguistic Variables: Lessons From Variable (ING). *Frontiers in artificial intelligence* 4: 648543.
- Ladefoged, Peter. 1996. *Elements of acoustic phonetics* : University of Chicago Press.
- Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques* : Wiley-Blackwell.
- Rogers, Catherine L. Merete M. Glasbrenner, Teresa M. DeMasi and Michelle Bianchi. 2013. Vowel inherent spectral change and the second-language learner. *Vowel inherent spectral change*: 231–259.
- Thomas, Erik R. 2011. *Sociophonetics: An introduction*. Basingstoke, Hampshire, New Yoek, NY: Palgrave Macmillan.
- van Rossum, Guido and Fred L. Drake. 2023. *Python*.

# A Fundamental Shift in NLP? Traditional Models vs. Neural Networks and Large Language Models – Observations in Student Bachelor's and Master's Theses

Robin Jegan

Lehrstuhl für Medieninformatik  
Otto-Friedrich-Universität Bamberg  
[robin.jegan@uni-bamberg.de](mailto:robin.jegan@uni-bamberg.de)

ChatGPT and Large Language Models (LLM) have transformed the Natural Language Processing landscape since the release of ChatGPT in November 2022. The capabilities are widely documented and, in many cases, demonstrably impressive. However, several problems, such as hallucinations in the generated texts, to name just one, persist and even with updates using larger models, or with the release of competing products such as Bing Chat or Bard, could not be solved adequately.

In this presentation, we look back on several bachelor's and master's thesis that have been written at the Chair of Media Informatics and show, that *modern techniques* (namely neural networks in a similar vein as LLMs) have their advantages, while traditional and *simpler techniques*, such as rule-based systems or support vector machines can match or outperform the results of newer models in certain scenarios.

The presentation is a brief preview of an in-progress dissertation dealing with the issues mentioned above, also written at the Chair of Media Informatics at the University of Bamberg. We will present an outlook for domain-specific use-cases and datasets, where state-of-the-art models will be compared to traditional models and advantages as well as disadvantages, depending on the use-case, will be discussed.

## References:

- I. Augenstein et al., "Factuality Challenges in the Era of Large Language Models," 2023, doi:10.48550/ARXIV.2310.05189.
- D. Aumiller, J. Fan, and M. Gertz, "On the State of German (Abstractive) Text Summarization," 2023, doi:10.48550/ARXIV.2301.07095.
- T. B. Brown et al., "Language Models are Few-Shot Learners," 2020, doi:10.48550/ARXIV.2005.14165.

## Deutsch-Chinesisches Lernerkorpus – DeChiLKO

Ming Liu

Professur Interkulturalität und Mehrsprachigkeit  
Georg-August Universität Göttingen  
[ming.liu@stud.uni-goettingen.de](mailto:ming.liu@stud.uni-goettingen.de)

In meiner Dissertation „Orthographiekompetenz chinesischer Deutschlernender – eine lernerkorpusbasierte Untersuchung“ wird das Chinesisch-deutsche Lerner-Korpus DeChiLKO erstellt. Auf META-LING 2023 möchte ich das Korpus und die aktuellen Ergebnisse vorstellen.

DeChiLKO besteht aus einem Prüfungskorpus und einem Erwerbskorpus. Das Prüfungskorpus enthält 195 Diktate von Germanistikstudierenden unter den PGG-Prüfungsbedingungen (Prüfungszentrum 2013). Das Erwerbskorpus enthält Diktatübungen von 5 chinesischen Germanistikstudierenden während ihrer ersten drei Semestern. Alle Lernertexte werden nach der Tokenisierung unter der Anwendung des Treetaggers in EXMARaLDA (Dulko) (Nolda 2019) maschinell getaggt. Als Zielhypothese wird ein Diktatlösungstext manuell eingefügt, der gleichfalls automatisch annotiert wird. Abweichungen zwischen Zielhypothese und Lernertext werden maschinell erkannt.

Der spezielle Fokus von DeChiLKO liegt darauf, dass die orthographischen Abweichungen an jeder Position bezüglich verschiedener orthographischer Dimensionen in den jeweiligen Fehlerebenen annotiert werden. Die Tokens werden nach Graphemen, Silbenstrukturen und morphematischen Strukturen segmentiert und mit vordefinierten Fehler-Tagsets markiert. Der gesamte Annotationsprozess erfolgt in EXMARaLDA (Dulko) (Hirschmann & Nolda 2019; Nolda 2019).

Nach einer Konvertierung in makedulko ist das gesamte DeChiLKO-Korpus (308 Texte mit 31.674 Tokens) als ANNIS-Korpus (Krause & Zeldes 2016) zugänglich und ermöglicht eine gründliche Korpussuche. Derzeit befindet sich meine Arbeit in der Auswertungsphase, die darauf abzielt, bedeutende Erkenntnisse zur Orthographiekompetenz chinesischer Deutschlernender zu liefern.

### Literatur:

- Hirschmann, Hagen & Andreas Nolda (2019). Dulko – auf dem Weg zu einem deutsch-ungarischen Lernerkorpus. In Neues vom heutigen Deutsch: Empirisch – methodisch – theoretisch. Institut für Deutsche Sprache: Jahrbuch 2018, Eichinger Ludwig & Albrecht Plewnia (Hgg.), 339–342. Berlin: De Gruyter.
- Krause, Thomas & Amir Zeldes (2016). ANNIS3: A new architecture for generic corpus

query and visualization. *Digital Scholarship in the Humanities* 31 (1): 118–139.

Nolda, Andreas (2019). Annotation von Lernerdaten mit EXMARaLDA (Dulko). <http://andreas.nolda.org> (19.11.2021).

Prüfungszentrum (2013). Prüfungsordnungen für das Germanistik-Grundstudium und -Hauptstudium im Hochschulwesen Chinas. Shanghai: Verlag für Fremdsprachenausbildung in Shanghai.

# From Byte to Babel: Large Language Models and the Tower of Linguistic

## Knowledge

Bastian Bunzeck

Computational Linguistics Group (CLAUSE)

Universität Bielefeld

[bastian.bunzeck@uni-bielefeld.de](mailto:bastian.bunzeck@uni-bielefeld.de)

The release and adoption of large language models like ChatGPT has generated significant attention in the scientific and public discourse. These models have surpassed traditional NLP methods in many areas while also proving practically useful. Furthermore, they lend themselves to all kinds of linguistic investigations. For example, both real and artificial neural networks exhibit similar activation patterns when processing language, written [Schrimpf et al., 2021, Hosseini et al., 2022] or spoken [Begu's et al., 2023], suggesting shared underlying mechanisms. Additionally, studies have revealed that different aspects of language acquisition like phonological effects in word learning [Jones and Brandt, 2020], the positive effects of child-directed speech on scaffolding [Huebner et al., 2021] or linguistic rule formation [Yedetore et al., 2023] are mirrored in neural models. Moreover, models like BERT have been shown to rediscover classical linguistic pipelines through their layers [Tenney et al., 2019], encompassing phonetics, morphology, syntax, and more. Interestingly, these findings are not entirely novel; as similar approaches were tested as early as the 1980s (Parallel Distributed Processing, [Rumelhart and McClelland, 1986]) and 1990s [Elman et al., 1996], culminating in the past-tense debate [Seidenberg and Plaut, 2014]. However, despite employing the same technologies, all of these approaches lack a unified theoretical paradigm. Therefore, a crucial question arises: What should large language models strive to be, and what linguistic assumptions underlie them? While some consider them to be ideal anti-Chomskyan, usage-based learners [Pannitto and Herbelot, 2022, Piantadosi, 2023], typical evaluation suites like BLiMP [Warstadt et al., 2020] routinely lie on Generative notions of grammaticality. Also, alternative approaches like construction grammar [Weissweiler et al., 2023] appear to slowly gain acceptance. To address these questions, the proposed presentation will provide an overview of the theoretical perspectives that underpin current research on large language models and the research questions that researchers aim to (dis)prove using these models. Additionally, it will introduce our own work, which seeks to bridge the gap between theoretical linguistics and computational models.

### References:

- [Begu's et al., 2023] Begu's, G., Zhou, A., and Zhao, T. C. (2023). Encoding of speech in convolutional layers and the brain stem based on language experience. *Scientific Reports*, 13(1):6480.
- [Elman et al., 1996] Elman, J. L., Bates, E. L., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. Neural Network Modeling and Connectionism. MIT Press, Cambridge, MA.

- [Hosseini et al., 2022] Hosseini, E. A., Schrimpf, M., Zhang, Y., Bowman, S., Zaslavsky, N., and Fedorenko, E. (2022). Artificial neural network language models align neurally and behaviorally with humans even after a developmentally realistic amount of training. Preprint, Neuroscience.
- [Huebner et al., 2021] Huebner, P. A., Sulem, E., Cynthia, F., and Roth, D. (2021). Baby-BERTa: Learning More Grammar With Small-Scale Child-Directed Language. In *Proceedings of the 25th Conference on Computational Natural Language Learning*, pages 624–646, Online. Association for Computational Linguistics.
- [Jones and Brandt, 2020] Jones, S. D. and Brandt, S. (2020). Density and Distinctiveness in Early Word Learning: Evidence From Neural Network Simulations. *Cognitive Science*, 44(1).
- [Pannitto and Herbelot, 2022] Pannitto, L. and Herbelot, A. (2022). Can Recurrent Neural Networks Validate Usage-Based Theories of Grammar Acquisition? *Frontiers in Psychology*, 13:741321.
- [Piantadosi, 2023] Piantadosi, S. T. (2023). Modern language models refute Chomsky’s approach to language.
- [Rumelhart and McClelland, 1986] Rumelhart, D. E. and McClelland, J. L. (1986). On Learning the Past Tenses of English Verbs. In *Parallel Distributed Processing*, volume 2, pages 535–551. MIT Press, Cambridge, MA.
- [Schrimpf et al., 2021] Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., and Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences*, 118(45):e2105646118.
- [Seidenberg and Plaut, 2014] Seidenberg, M. S. and Plaut, D. C. (2014). Quasiregularity and Its Discontents: The Legacy of the Past Tense Debate. *Cognitive Science*, 38(6):1190–1228.
- [Tenney et al., 2019] Tenney, I., Das, D., and Pavlick, E. (2019). BERT Rediscovered the Classical NLP Pipeline. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4593–4601, Florence, Italy. Association for Computational Linguistics.
- [Warstadt et al., 2020] Warstadt, A., Parrish, A., Liu, H., Mohananey, A., Peng, W., Wang, S.-F., and Bowman, S. R. (2020). BLiMP: The Benchmark of Linguistic Minimal Pairs for English. *Transactions of the Association for Computational Linguistics*, 8:377–392.
- [Weissweiler et al., 2023] Weissweiler, L., He, T., Otani, N., Mortensen, D. R., Levin, L., and Schütze, H. (2023). Construction Grammar Provides Unique Insight into Neural Language Models. In *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, pages 85–95.
- [Yedetore et al., 2023] Yedetore, A., Linzen, T., Frank, R., and McCoy, R. T. (2023). How poor is the stimulus? Evaluating hierarchical generalization in neural networks trained on child-directed speech.