Genderinklusive Movierung im Deutschen

Visuelle Formvariation und lexikalische Repräsentation

Dominic Schmitz

Heinrich Heine Universität Düsseldorf

Femininmovierung in Grammatik und Diskurs / Feminine Gender Marking in Grammar and Discourse
10. & 11. September 2026 · Københavns Universitet

Genderinklusive Movierung

- Bisherige Studien legen nahe, dass genderinklusiv movierte Formen
  - keinen männlichen **Bias** tragen (z. B. Körner et al. 2022)
  - unterschiedliche **Akzeptanz** in der Sprachgemeinschaft finden (z. B. Jäckle 2022)
  - nicht weniger **gut lesbar** sind als andere Formen
    - Stern (Friedrich et al. 2021; Pabst & Kollmayer 2023; Friedrich et al. 2024; Zacharski & Ferstl 2024; Haan 2025; Zacharski, Kruppa & Ferstl 2025; Bach et al. 2026; Decock et al. 2026)
    - Doppelpunkt (Bach et al. 2026)
    - Unterstrich (?)
- Auch einige wenige computationale Studien zu genderinklusiver Movierung mit Stern liegen bereits vor und komplementieren die vor allem psycholinguistischen Ergebnisse (z. B. Schmitz 2026)

Genderinklusive Form

- Genderinklusive Movierung erzeugt eine visuell charakteristische Wortform

Florist Floristin Florist*in Florist:in Florist_in

- Asterisk, Doppelpunkt und Unterstrich sind besonders **salient**
- Doch sind sie gleichermaßen salient und damit auch potenziell gleichermaßen das Lesen beeinflussend?
- Die derzeitige Studienlage lässt keine Antwort zu

Genderinklusive Form aus computationaler Sicht

- Die bisherigen computationalen Studien verwenden **orthographische Trigramme** als Repräsentationsform

```
#fl flo lor ori ris ist st* t*i *in in#
#fl flo lor ori ris ist st: t:i :in in#
#fl flo lor ori ris ist st_ t_i _in in#
```

- Da *, : und _ an identischer Stelle vorkommen, sind die drei Symbole austauschbar – sie sind **distributionell äquivalent**
- Ein computationales Modell, das nur * kennt, kommt daher zu den exakt gleichen Ergebnissen wie ein Modell, das nur : oder _ kennt
- Die tatsächliche Form – und damit auch potenzielle Unterschiede der Salienz – wird somit schlichtweg nicht berücksichtigt

Ziele der aktuellen Studie

Ziel 1

Update des bisherigen computationalen Ansatzes durch Mittel, die Unterschiede zwischen *, : und _ ermöglichen

Ziel 2

Überprüfung des aktualisierten computationalen Ansatzes durch Modellierung erhobener Lesezeitdaten für Sternformen

Computationaler Ansatz

Linear Discriminative Learning

- Das **Discriminative Lexicon Model** beruht auf der Annahme, dass Sprache durch Erfahrung modelliert wird (Baayen et al. 2019; Heitmeier et al. 2026)
  - Etabliertes psychologisches Lernmodell nach Rescorla & Wagner (1972)
- Diese Annahme ist computational durch **Linear Discriminative Learning** implementiert (Baayen et al. 2019; Heitmeier et al. 2026)
- Hier werden **Assoziationsgewichte** zwischen der Repräsentation von Form und Semantik erstellt
  - Kommen eine bestimmte Form und Semantik gemeinsam vor, wird ihr Assoziationsgewicht erhöht
  - Kommt eine bestimmte Form ohne eine bestimmte Semantik vor (oder vice versa), wird ihr Assoziationsgewicht erniedrigt

Linear Discriminative Learning

- Form und Semantik werden hierbei in Form von **Matrizen** repräsentiert

$$C = \begin{array}{c} \\ \textit{Maus} \\ \textit{Haus} \end{array} \begin{array}{c} \begin{array}{cccccc} \#ma & mau & aus & us\# & \#ha & hau & \dots \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 1 & 1 & 1 & \dots \end{pmatrix} \end{array}$$

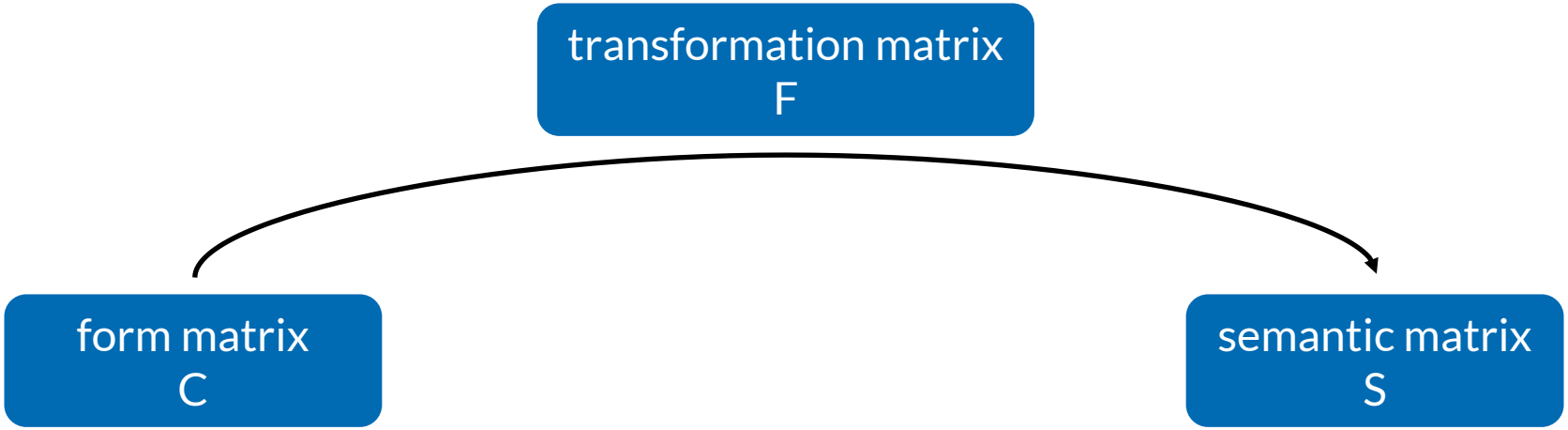
$$S = \begin{array}{c} \\ \textit{Maus} \\ \textit{Haus} \end{array} \begin{array}{c} \begin{array}{ccccc} S1 & S2 & S3 & \dots \end{array} \\ \begin{pmatrix} 29 & 1 & -1 & \dots \\ -10 & 15 & -10 & \dots \end{pmatrix} \end{array}$$

Linear Discriminative Learning

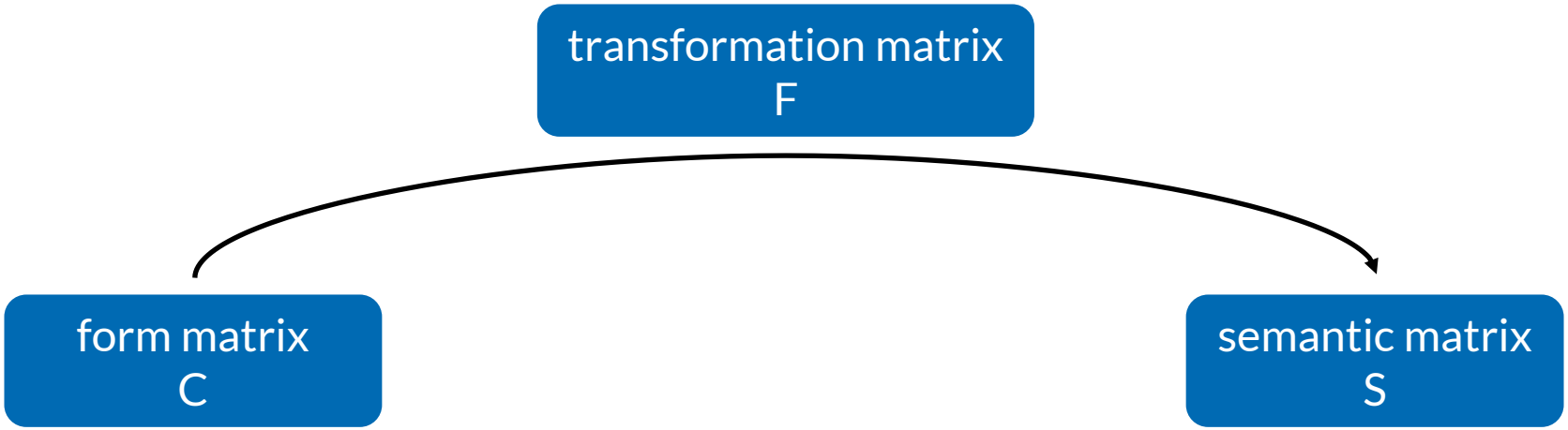
- Wird von der Form-Matrix ausgegangen und ist die Semantik-Matrix das Ziel, so wird die Sprachverarbeitung simuliert

  → gegeben sei eine Form c in C, wie lautet ihre Bedeutung s in S?

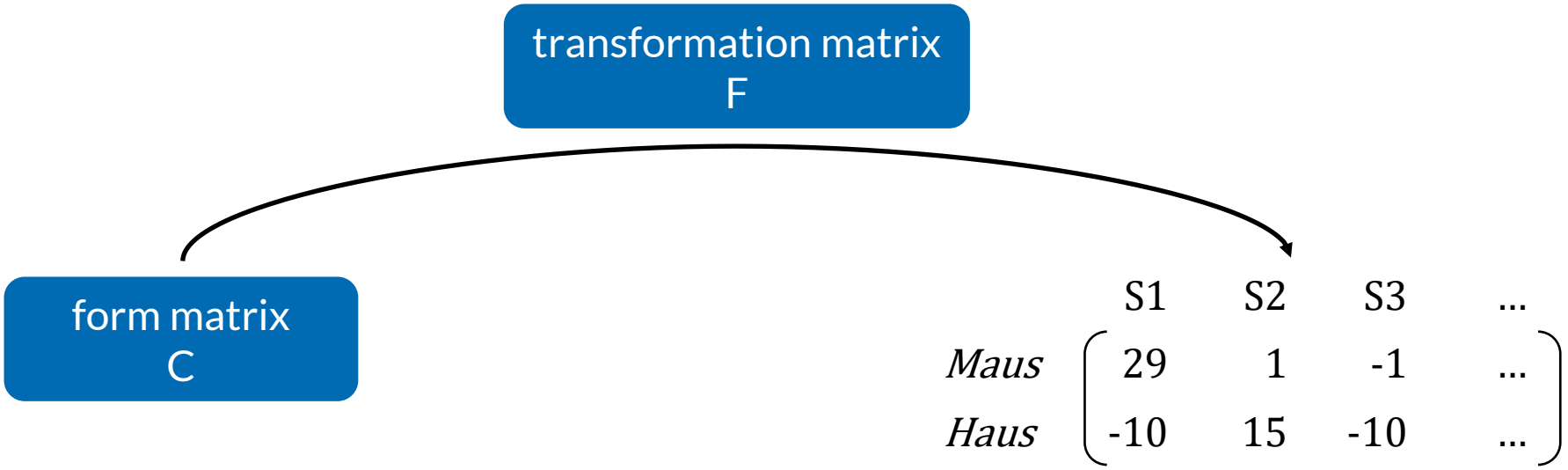
- Durch Matrixmultiplikation, d.h. anhand einer Transformationsmatrix, kann die Form-Matrix zur Semantik-Matrix umgeformt werden



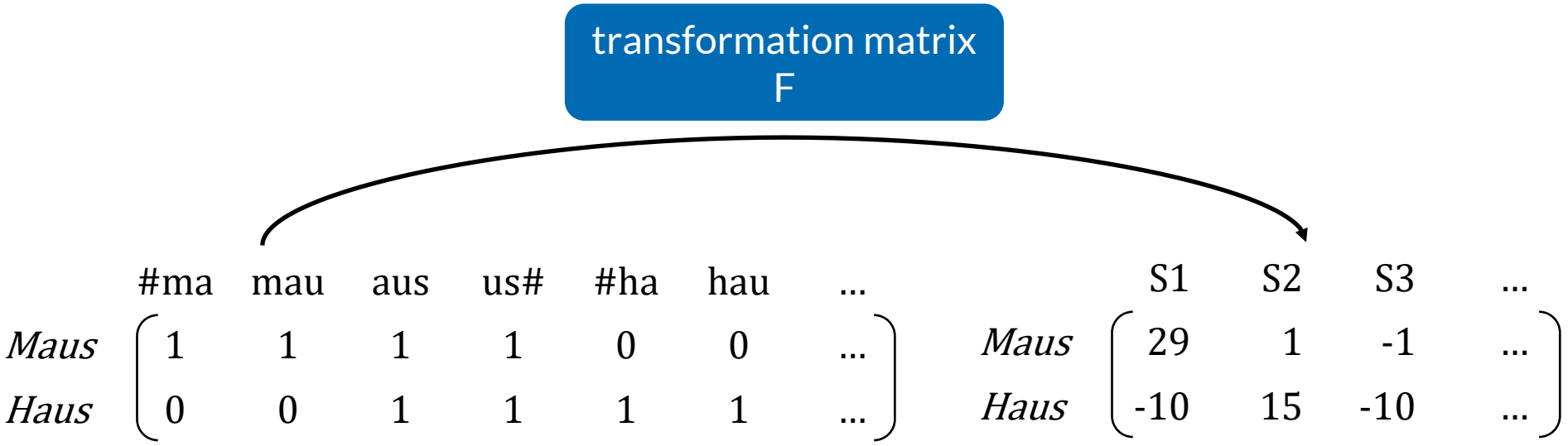
Linear Discriminative Learning



Linear Discriminative Learning



Linear Discriminative Learning



Linear Discriminative Learning

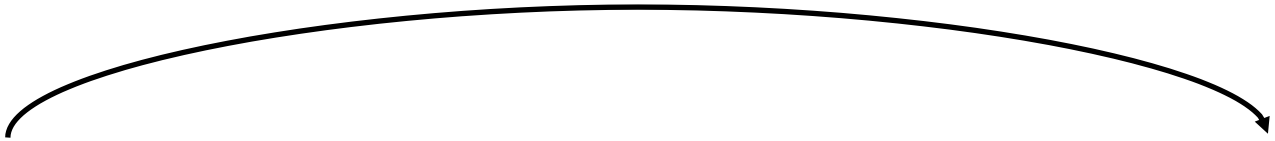
	S1	S2	S3
#ma	11.33	-2.17	1.33
mau	11.33	-2.17	1.33
aus	3.17	2.67	-1.83
us#	3.17	2.67	-1.83
#ha	-8.17	4.83	-3.17
hau	-8.17	4.83	-3.17

	#ma	mau	aus	us#	#ha	hau	...
Maus	1	1	1	1	0	0	...
Haus	0	0	1	1	1	1	...

	S1	S2	S3	...
Maus	29	1	-1	...
Haus	-10	15	-10	...

Linear Discriminative Learning

	S1	S2	S3
#ma	11.33	-2.17	1.33
mau	11.33	-2.17	1.33
aus	3.17	2.67	-1.83
us#	3.17	2.67	-1.83
#ha	-8.17	4.83	-3.17
hau	-8.17	4.83	-3.17



	#ma	mau	aus	us#	#ha	hau	...
Maus	1	1	1	1	0	0	...
Haus	0	0	1	1	1	1	...

	S1	S2	S3	...
Maus	28.9	1.1	-1.2	...
Haus	-9.1	14.2	-9.9	...

Annäherung an *S*

Linear Discriminative Learning

	S1	S2	S3
#ma	11.33	-2.17	1.33
mau	11.33	-2.17	1.33
aus	3.17	2.67	-1.83
us#	3.17	2.67	-1.83
#ha	-8.17	4.83	-3.17
hau	-8.17	4.83	-3.17

	#ma	mau	aus	us#	#ha	hau	…
Maus	1	1	1	1	0	0	…
Haus	0	0	1	1	1	1	…

	S1	S2	S3	…
Maus	28.9	1.1	-1.2	…
Haus	-9.1	14.2	-9.9	…

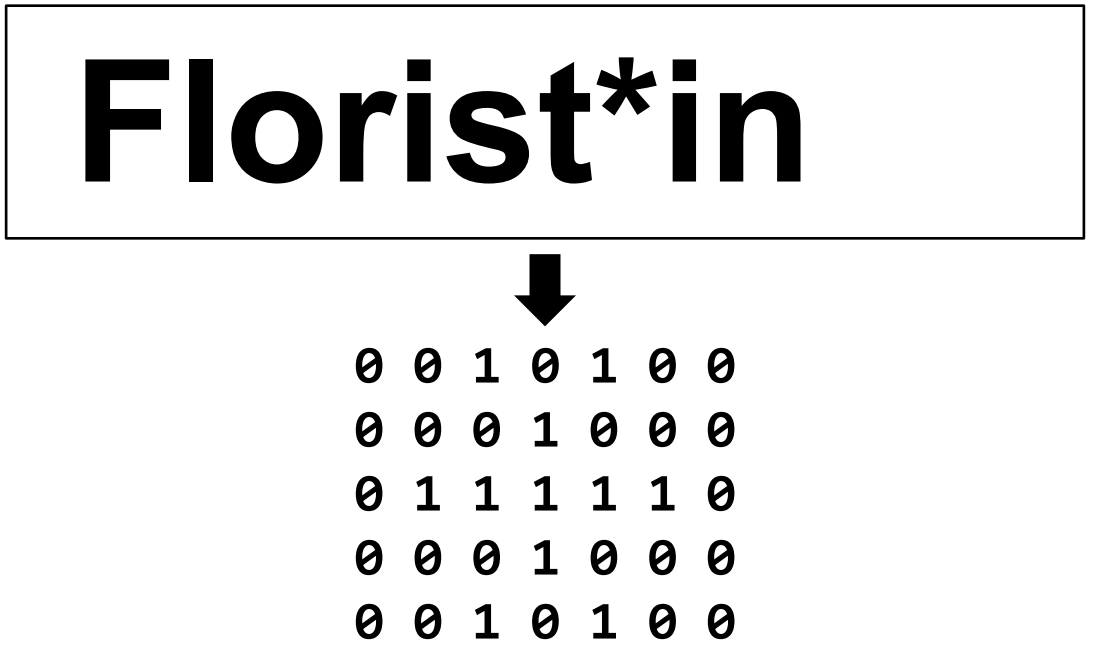
das Problem: so sind *, : und _ austauschbar

Pixel statt Trigramme

Zeichenfolge als Ausgangspunkt

Florist*in

gleiche Schriftart, Schriftgröße, Bildfläche, Ausrichtung für alle Wörter



Jeder Bildpunkt wird durch einen numerischen Wert repräsentiert

Pixelvektor

```
[0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, …]
```

Pixel statt Trigramme

$$C = \begin{array}{c} \\ \textit{Maus} \\ \textit{Haus} \end{array} \begin{array}{c} \begin{array}{cccccc} \#ma & mau & aus & us\# & \#ha & hau & \dots \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 1 & 1 & 1 & \dots \end{pmatrix} \end{array}$$

$$C = \begin{array}{c} \\ \textit{Maus} \\ \textit{Haus} \end{array} \begin{array}{c} \begin{array}{cccccc} P1 & P2 & P3 & P4 & P5 & P6 & \dots \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 & \dots \\ 1 & 1 & 0 & 1 & 1 & 0 & \dots \end{pmatrix} \end{array}$$

Maße

- Aus approximierter C und S Matrix lässt sich eine Reihe von **Maßen** zu Bedeutung, Form, Verarbeitung und Produktion berechnen (z. B. Chuang et al. 2021; Schmitz et al. 2021)
- Die hier wichtigen Maße
  - **Verständnisqualität**

    Wie nah am ursprünglichen semantischen ist der verstandene Vektor?

    → mathematisch: Korrelation
  - **Form-Nachbarschaftsdichte**

    Wie nah liegen die zehn nächsten Nachbarn an der produzierten Form?

    → mathematisch: durchschnittliche Korrelation

Lesezeitdaten

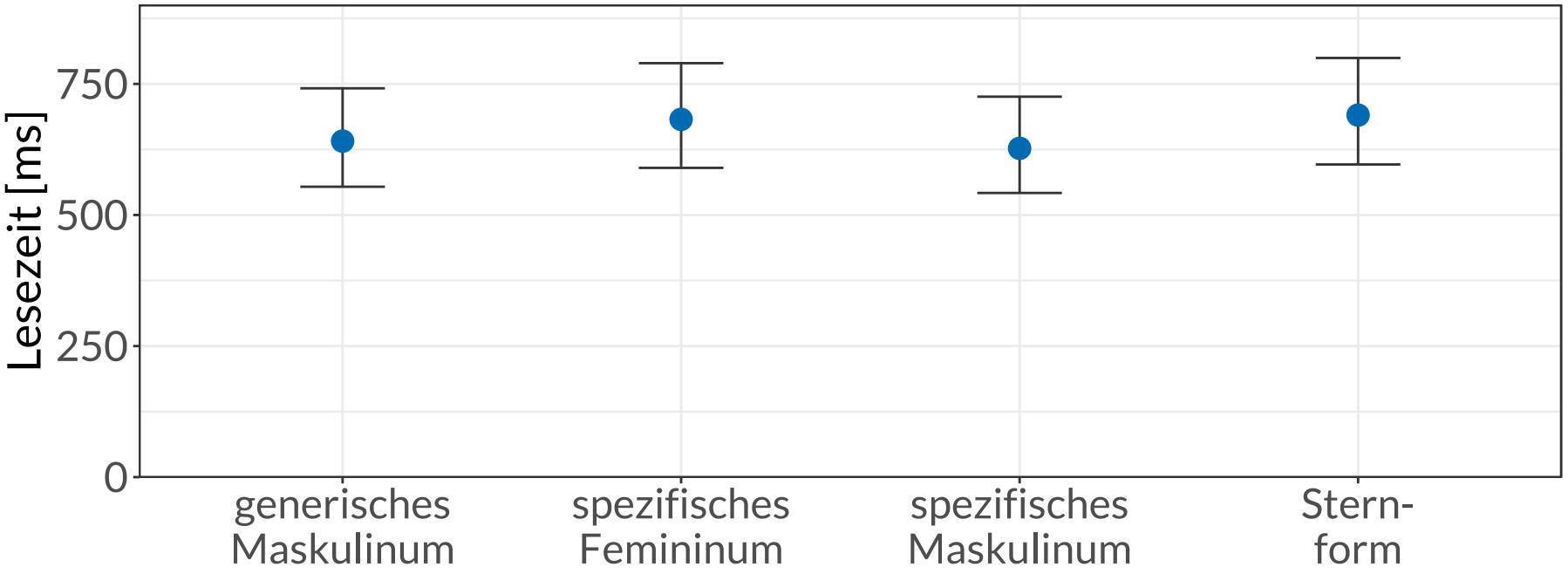
Self-Paced-Reading-Studie: Design

- Die Daten, die wir heute analysieren, stammen aus einer Self-Paced-Reading-Studie (Haan 2025)
  - 119 Versuchspersonen
  - 24 Personenbezeichnungen
  - wortweises Moving-Window-Design
  - abhängige Variable: Lesezeit am Zielwort
- Beispiel für alle 4 BEDINGUNGen:

	Mädchen	Floristinnen
Viele	Kinder	lieben Blumen. Sie möchten als Florist*innen arbeiten, wenn sie erwachsen sind.
	Kinder	Floristen
	Jungen	Floristen

Self-Paced-Reading-Studie: Ergebnis

- BEDINGUNG zeigt einen Effekt
  - Spezifisches Maskulinum schneller als spezifisches Femininum und als Sternform
  - Kein Unterschied zwischen generischem Maskulinum und Sternform



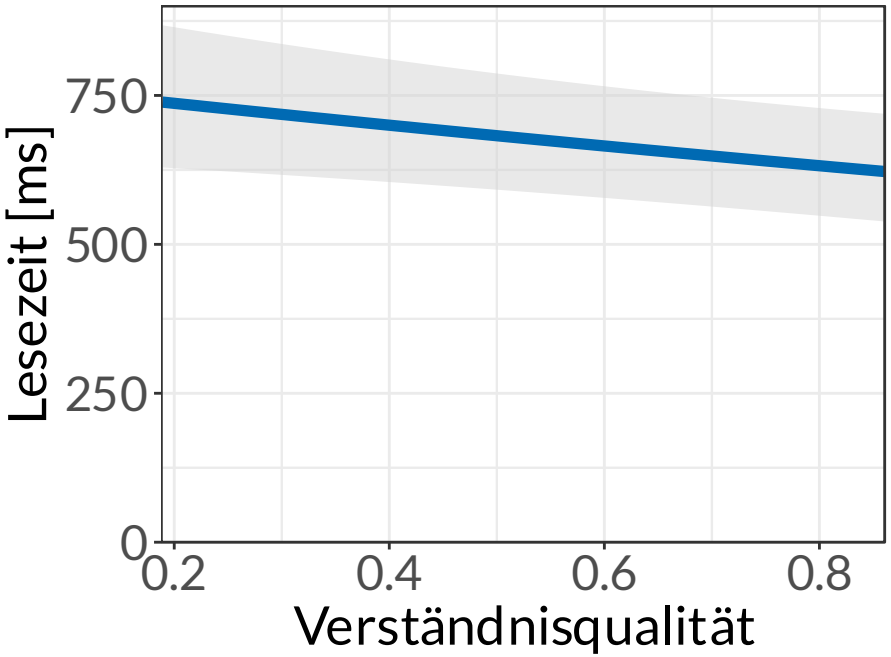
Was offen bleibt

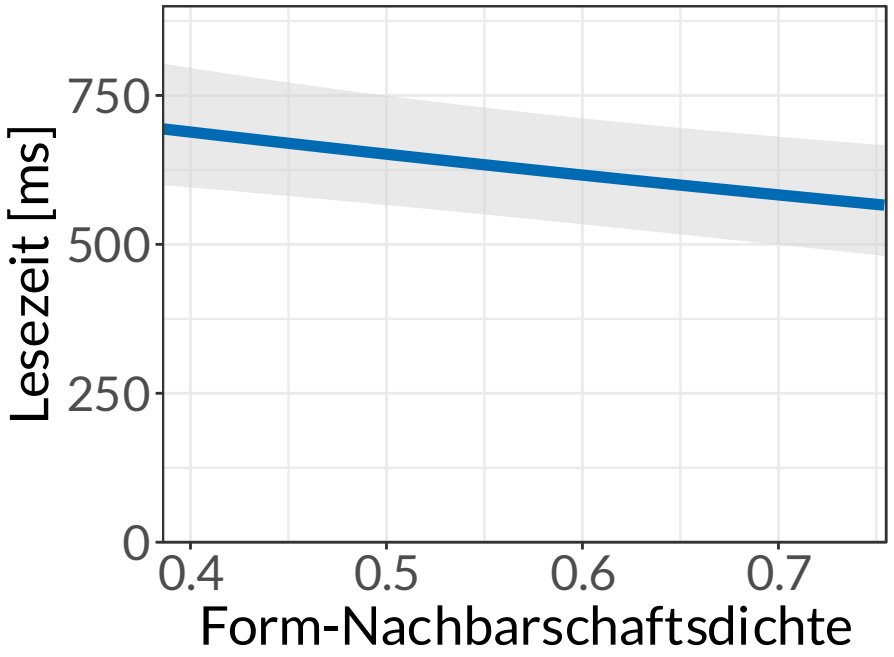
- BEDINGUNG bündelt verschiedene Arten von Information
  - sichtbare Form inkl. Unterschied zwischen *, : und _
  - Wortlänge
  - Frequenz
  - intendierte Bedeutung
  - …
- Können wir für die Lesezeit relevante lexikalische Eigenschaften beschreiben?

  → genau hier setzen die LDL-Maße an

Lesezeiten ~ LDL-Maße

- Statt BEDINGUNG werden LDL-Maße in der Analyse verwendet
- Dabei kommt heraus: **Verständnisqualität** und **Form-Nachbarschaftsdichte** sagen Lesezeiten vorher

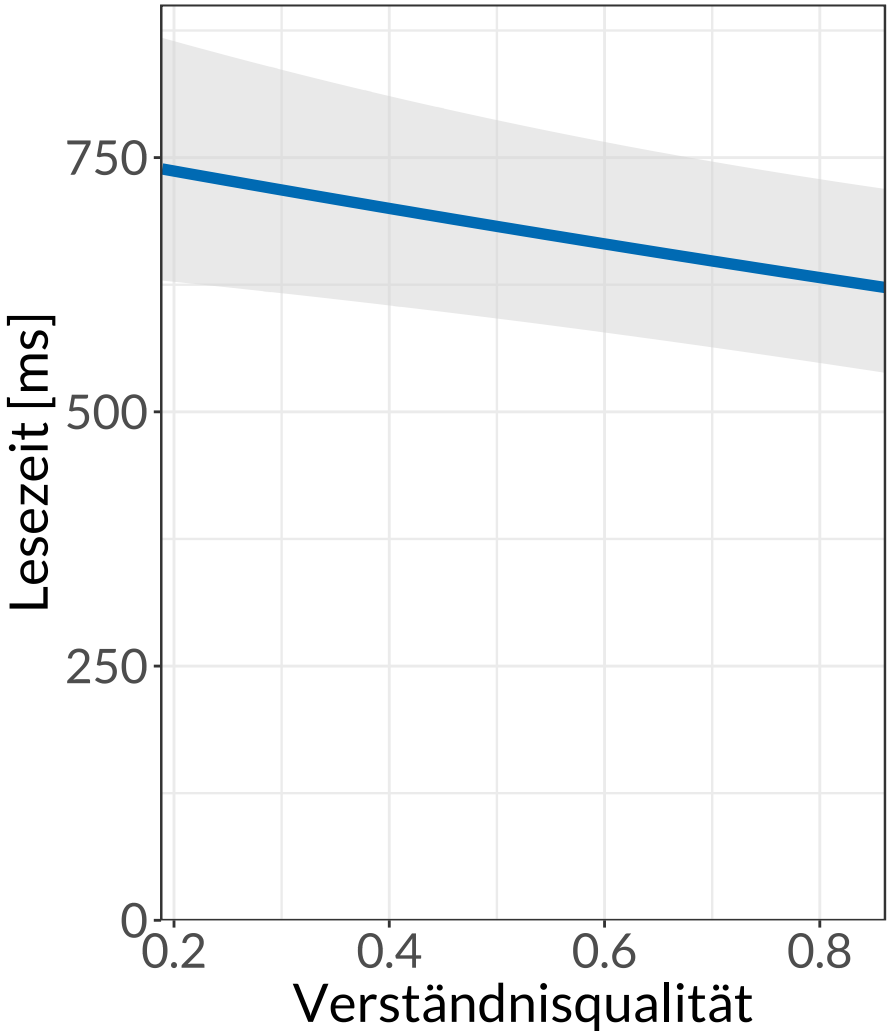




Verständnisqualität

Sternform Femininum > Maskulinum

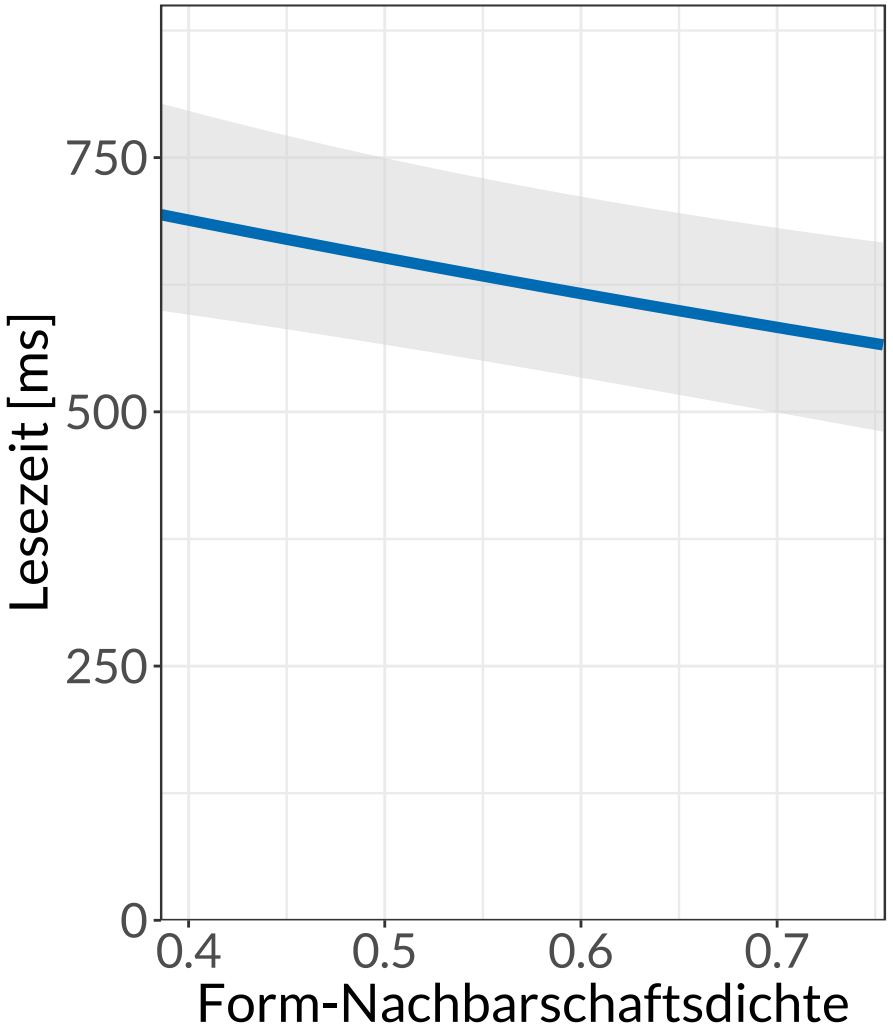
- **Stern** und **Femininum** sind einfacher zu verstehen, weil sie distinkt in ihrer Bedeutung sind
- **Maskulina** sind schwieriger zu verstehen, weil ihre Bedeutungen so ähnlich sind
- **Effekt**: Wer einfacher zu verstehen ist, wird schneller gelesen



Form-Nachbarschaftsdichte

Maskulinum > Sternform Femininum

- **Stern** und **Femininum** haben weniger dichte Nachbarschaften, da es weniger Wörter mit ihrer Form gibt
- **Maskulina** haben dichtere Nachbarschaften, da ihre Formen identisch sind
- **Effekt**: Wer mehr Nachbarn hat, wird schneller gelesen



Hä?

- Wie passen diese Ergebnisse mit denen der Masterarbeit zusammen?
- Zur Erinnerung
  - Spezifisches Maskulinum schneller als spezifisches Femininum und als Sternform
  - Kein Unterschied zwischen generischem Maskulinum und Sternform
- Der Effekt von **Form-Nachbarschaftsdichte** ist stärker als der von **Verständnisqualität**
- Da diese „gegeneinander“ arbeiten, gewinnt der stärkere Effekt: Maskulina sind schneller zu lesen
- BEDINGUNG verliert im LDL-Modell sämtliche Erklärkraft

Zusammenfassung & Diskussion

Zusammenfassung: Ziel 1

Ziel 1

Update des bisherigen computationalen Ansatzes durch Mittel, die Unterschiede zwischen *, : und _ ermöglichen

- Pixelbasierte Vektoren unterscheiden zwischen Asterisk, Doppelpunkt und Unterstrich
- Inwiefern diese Unterschiede behaviorale Konsequenzen – und damit auch Konsequenzen auf entsprechende computationale Analysen – haben, bleibt derzeit aufgrund fehlender experimenteller Daten offen
- Einzelne LDL-Maße zeigen klare Unterschiede zwischen den Formen

Zusammenfassung: Ziel 2

Ziel 2

Überprüfung des aktualisierten computationalen Ansatzes durch Modellierung erhobener Lesezeitdaten für Sternformen

- Pixelbasierte Vektoren gewähren einen detaillierteren Blick auf Unterschiede in Lesezeitdaten
- Statt kategorialer Unterschiede ermöglichen sie Rückschlüsse darauf, was im Detail zu Lesezeitunterschieden führt
- Hier
  - **Verständnisqualität**: je besser, desto schneller
  - **Form-Nachbarschaftsdichte**: je dichter, desto schneller

Diskussion

- Experimentelle Daten zum Lesen von Formen jenseits des Asterisks werden dringend benötigt
- Liegen diese vor, kann der hier vorgestellte computationale Ansatz auf diese sinnvoll erweitert werden
- Solange diese nicht vorliegen, können zwar Unterschiede in Pixelvektoren und entsprechenden Maßen gefunden werden, inwiefern diese jedoch aussagekräftig sind, bleibt offen
- Nichtsdestotrotz ist die kontinuierliche Formrepräsentation durch Pixelvektoren näher an dem, was Menschen beim Lesen tatsächlich aufnehmen, und damit ein Schritt in Richtung naturalistischer Modellierung

Vielen Dank!

Literatur 1/2

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, *2019*, 4895891. https://doi.org/10.1155/2019/4895891

Bach, K. M., Pickal, A. J., Wachholz, V., Richters, C., Murböck, J., Brandl, L., Stadler, M., & Sailer, M. (2026). *Gendergerecht oder Gendergaga? Eine Metaanalyse über den Einfluss von geschlechtergerechten Formulierungen auf das Lesen*. PsyArXiv. https://doi.org/10.31234/osf.io/9qzvs_v2

Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & Baayen, R. H. (2021). The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods*, *53*(3), 945–976. https://doi.org/10.3758/s13428-020-01356-w

Decock, S., Zacharski, L., Boone, G., & Ferstl, E. C. (2026). Comprehensibility of gender-fair language among foreign language learners of German: An experimental study. *Frontiers in Language Sciences*, *5*. https://doi.org/10.3389/flang.2026.1806497

Friedrich, M. C. G., Drößler, V., Oberlehberg, N., & Heise, E. (2021). The Influence of the Gender Asterisk (“Gendersternchen”) on Comprehensibility and Interest. *Frontiers in Psychology*, *12*, 5934. https://doi.org/10.3389/FPSYG.2021.760062/BIBTEX

Friedrich, M. C. G., Gajewski, S., Hagenberg, K., Wenz, C., & Heise, E. (2024). Does the gender asterisk (“Gendersternchen”) as a special form of gender-fair language impair comprehensibility? *Discourse Processes*, *61*(9), 439–461. https://doi.org/10.1080/0163853X.2024.2362027

Haan, A.-S. (2025). *Gender-fair language processing: A self-paced reading study on the German gender star* [Master]. Heinrich Heine University Düsseldorf.

Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2026). The discriminative lexicon: Theory, implementation in the Julia package JudiLing, and applications. Cambridge University Press.

Literatur 2/2

Jäckle, S. (2022). Per aspera ad astra – Eine politikwissenschaftliche Analyse der Akzeptanz des Gendersterns in der deutschen Bevölkerung auf Basis einer Online-Umfrage. Politische Vierteljahresschrift, 63(3), 469–497. https://doi.org/10.1007/s11615-022-00380-z

Körner, A., Abraham, B., Rummer, R., & Strack, F. (2022). Gender representations elicited by the gender star form. Journal of Language and Social Psychology, 41(5), 553–571. https://doi.org/10.1177/0261927X221080181

Pabst, L. M., & Kollmayer, M. (2023). How to make a difference: The impact of gender-fair language on text comprehensibility amongst adults with and without an academic background. Frontiers in Psychology, 14, 1234860. https://doi.org/10.3389/FPSYG.2023.1234860

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Hrsg.), *Classical conditioning II: Current research and theory* (S. 64–99). Appleton-Century-Crofts.

Schmitz, D., Plag, I., Baer-Henney, D., & Stein, S. D. (2021). Durational differences of word-final /s/ emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.680889

Schmitz, D. (2026). Die Repräsentation des Gendersternchens im mentalen Lexikon. Perspektiven der Movierungsforschung.

Zacharski, L., Kruppa, A., & Ferstl, E. C. (2025). The Readability of the Non-Binary Gender Star in German: Evidence From a Lexical Decision Task. Social Psychological Bulletin, 20, 1–30. https://doi.org/10.32872/spb.13719

Zielwörter

- **Weiblicher Bias**

  Schneider, Kindergärtner, Florist, Friseur, Flugbegleiter, Balletttänzer, Maskenbildner, Wahrsager

- **Kein Bias**

  Künstler, Politiker, Journalist, Fotograf, Entdecker, Arzt, Schriftsteller, Zoologe

- **Männlicher Bias**

  Rennfahrer, Astronaut, Automechaniker, Bauarbeiter, Detektiv, Fußballspieler, Programmierer, Polizist