

# Die semantische Repräsentation des generischen Maskulinums

Dominic Schmitz<sup>1</sup>, Viktoria Schneider<sup>1</sup>, Janina Esser<sup>2</sup>

<sup>1</sup>Heinrich-Heine-Universität Düsseldorf, <sup>2</sup>Diversität in der Linguistik e.V.

45. Jahrestagung der DGfS – AG10

Geschlechterbewusste Sprache: Routinisierte Gebrauch und kreative Neuförmungen

# Generische Maskulina im Deutschen

## Generische Maskulina im Deutschen

Wortform	Sexus	Genus	Numerus
<i>Lehrer</i>	männlich	Maskulinum	Singular
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerin</i>	weiblich	Femininum	
<i>Lehrer</i>	männlich	Maskulinum	Plural
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerinnen</i>	weiblich	Femininum	

## Generische Maskulina im Deutschen

Wortform	Sexus	Genus	Numerus
<i>Lehrer</i>	männlich	Maskulinum	Singular
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerin</i>	weiblich	Femininum	
<i>Lehrer</i>	männlich	Maskulinum	Plural
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerinnen</i>	weiblich	Femininum	

# Generische Maskulina im Deutschen

Rollenbezeichnungs-  
paradigma

Wortform	Sexus	Genus	Numerus
<i>Lehrer</i>	männlich	Maskulinum	Singular
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerin</i>	weiblich	Femininum	
<i>Lehrer</i>	männlich	Maskulinum	Plural
<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
<i>Lehrerinnen</i>	weiblich	Femininum	

# Generische Maskulina im Deutschen

	Wortform	Sexus	Genus	Numerus
Rollenbezeichnungs- paradigma	<i>Lehrer</i>	männlich	Maskulinum	Singular
	<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
	<i>Lehrerin</i>	weiblich	Femininum	
	<i>Lehrer</i>	männlich	Maskulinum	Plural
	<i>Lehrer</i>	männlich oder weiblich	Maskulinum	
	<i>Lehrerinnen</i>	weiblich	Femininum	

- Gemeinhin wird generischen Maskulina nachgesagt, dass sie in ihrer Bedeutung geschlechtsneutral seien

# Bisherige Forschung

## Bisherige Forschung

- Zweifel an der geschlechtsneutralen Bedeutung finden sich allerdings in den Ergebnissen bisheriger Forschung

## Bisherige Forschung

- Zweifel an der geschlechtsneutralen Bedeutung finden sich allerdings in den Ergebnissen bisheriger Forschung
- Die überwiegende Mehrheit bestehender Studien zum Thema kommt zu einem gemeinsamen Schluss
  - generische Maskulina sind nicht geschlechtsneutral, sondern zeigen einen **männlichen Bias** (z. B. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)

## Bisherige Forschung

- Zweifel an der geschlechtsneutralen Bedeutung finden sich allerdings in den Ergebnissen bisheriger Forschung
- Die überwiegende Mehrheit bestehender Studien zum Thema kommt zu einem gemeinsamen Schluss
  - generische Maskulina sind nicht geschlechtsneutral, sondern zeigen einen **männlichen Bias** (z. B. Demarmels, 2017; Garnham et al., 2012; Gyga et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)
- Obwohl generische Maskulina also durchaus geschlechtsneutral intendiert sein können...

## Bisherige Forschung

- Zweifel an der geschlechtsneutralen Bedeutung finden sich allerdings in den Ergebnissen bisheriger Forschung
- Die überwiegende Mehrheit bestehender Studien zum Thema kommt zu einem gemeinsamen Schluss
  - generische Maskulina sind nicht geschlechtsneutral, sondern zeigen einen **männlichen Bias** (z. B. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)
- Obwohl generische Maskulina also durchaus geschlechtsneutral intendiert sein können...
- ...kommt diese Geschlechtsneutralität im Sprachverständnis nicht (völlig) an

## Bisherige Forschung

- Zweifel an der geschlechtsneutralen Bedeutung finden sich allerdings in den Ergebnissen bisheriger Forschung
- Die überwiegende Mehrheit bestehender Studien zum Thema kommt zu einem gemeinsamen Schluss
  - generische Maskulina sind nicht geschlechtsneutral, sondern zeigen einen **männlichen Bias** (z. B. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)
- Obwohl generische Maskulina also durchaus geschlechtsneutral intendiert sein können...
- ...kommt diese Geschlechtsneutralität im Sprachverständnis nicht (völlig) an
- Stattdessen findet eine Verarbeitung statt, aus der eine männliche Überrepräsentation resultiert

# Probleme bisheriger Forschung

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenziellen Einflüsse von Stereotypen.

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenziellen Einflüsse von Stereotypen.

## Problem 2: Datengrundlage

Viele Studien nutzen Daten, die speziell für ihre jeweiligen Untersuchungen erhoben wurden.

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenziellen Einflüsse von Stereotypen.

## Problem 2: Datengrundlage

Viele Studien nutzen Daten, die speziell für ihre jeweiligen Untersuchungen erhoben wurden.

## Problem 3: Semantische Repräsentationen

Viele Studien liefern einen Bias, aber keine Erklärung für dessen semantische Grundlage.

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenzielle Einflüsse von Stereotypen.

→ Stereotypizität als Kovariate

## Problem 2: Datengrundlage

Viele Studien nutzen Daten, die speziell für ihre jeweiligen Untersuchungen erhoben wurden.

## Problem 3: Semantische Repräsentationen

Viele Studien liefern einen Bias, aber keine Erklärung für dessen semantische Grundlage.

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenzielle Einflüsse von Stereotypen.

→ Stereotypizität als Kovariate

## Problem 2: Datengrundlage

Viele Studien nutzen Daten, die speziell für ihre jeweiligen Untersuchungen erhoben wurden.

→ Korpusdaten als Datengrundlage

## Problem 3: Semantische Repräsentationen

Viele Studien liefern einen Bias, aber keine Erklärung für dessen semantische Grundlage.

# Probleme bisheriger Forschung

## Problem 1: Stereotype

Viele, insbesondere ältere, Studien zum Thema berücksichtigen keine potenzielle Einflüsse von Stereotypen.

→ Stereotypizität als Kovariate

## Problem 2: Datengrundlage

Viele Studien nutzen Daten, die speziell für ihre jeweiligen Untersuchungen erhoben wurden.

→ Korpusdaten als Datengrundlage

## Problem 3: Semantische Repräsentationen

Viele Studien liefern einen Bias, aber keine Erklärung für dessen semantische Grundlage.

→ Discriminative Learning für semantische Messwerte

# Forschungsfragen

# Forschungsfragen

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

# Forschungsfragen

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

## Frage 2

Falls ja, inwiefern unterscheidet sich die Semantik generischer Maskulina von der Semantik spezifischer Maskulina und Feminina?



# Methode

Discriminative Learning



# Grundidee

## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)

## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt

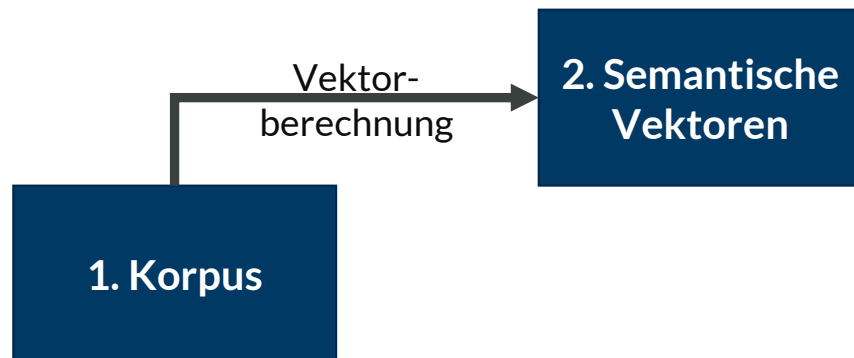
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt

1. Korpus

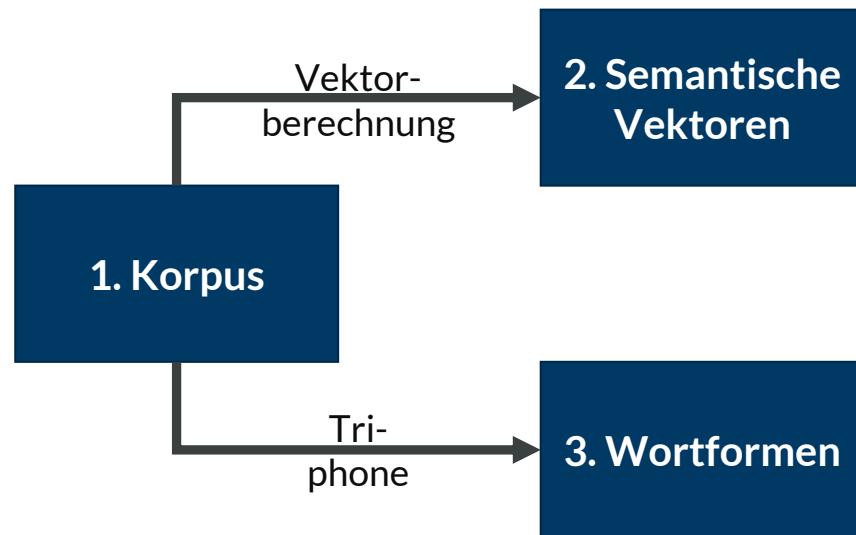
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



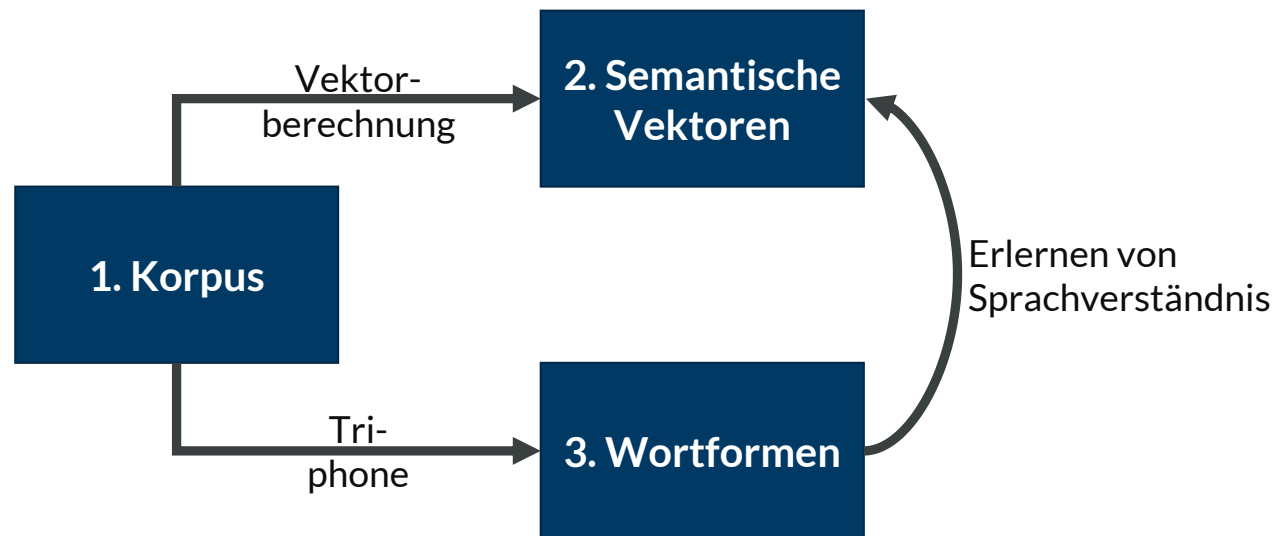
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt





# Korpus

# Korpus

- Das genutzte Korpus wurde anhand des Subkorpus „News“ der Leipzig Corpora Collection (Goldhahn et al., 2012) erstellt
  - 49.044.960 Wortformen
  - 30.000 Sätze mit Zielwörtern
  - 800.000 Sätze ohne Zielwörter

# Korpus

- Das genutzte Korpus wurde anhand des Subkorpus „News“ der Leipzig Corpora Collection (Goldhahn et al., 2012) erstellt
  - 49.044.960 Wortformen
  - 30.000 Sätze mit Zielwörtern
  - 800.000 Sätze ohne Zielwörter
- Zielwörter wurden aus einer vorherigen Studie übernommen, in welcher es um die Stereotypizität von Rollenbezeichnungen ging (Gabriel et al., 2008)

# Korpus

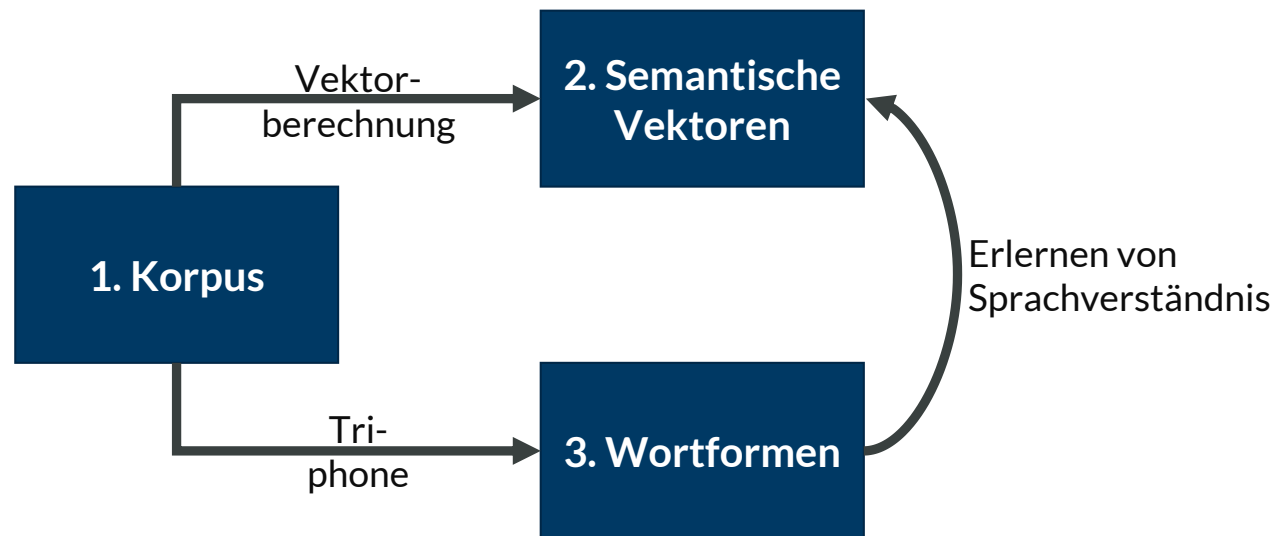
- Das genutzte Korpus wurde anhand des Subkorpus „News“ der Leipzig Corpora Collection (Goldhahn et al., 2012) erstellt
  - 49.044.960 Wortformen
  - 30.000 Sätze mit Zielwörtern
  - 800.000 Sätze ohne Zielwörter
- Zielwörter wurden aus einer vorherigen Studie übernommen, in welcher es um die Stereotypizität von Rollenbezeichnungen ging (Gabriel et al., 2008)
- Insgesamt 113 Rollenbezeichnungspadigmen wurden übernommen

# Korpus

- Das genutzte Korpus wurde anhand des Subkorpus „News“ der Leipzig Corpora Collection (Goldhahn et al., 2012) erstellt
  - 49.044.960 Wortformen
  - 30.000 Sätze mit Zielwörtern
  - 800.000 Sätze ohne Zielwörter
- Zielwörter wurden aus einer vorherigen Studie übernommen, in welcher es um die Stereotypizität von Rollenbezeichnungen ging (Gabriel et al., 2008)
- Insgesamt 113 Rollenbezeichnungspadigmen wurden übernommen
- Im Korpus sind keine „neuen“ Formen (Sternchenform, Binnen-I, etc.) enthalten

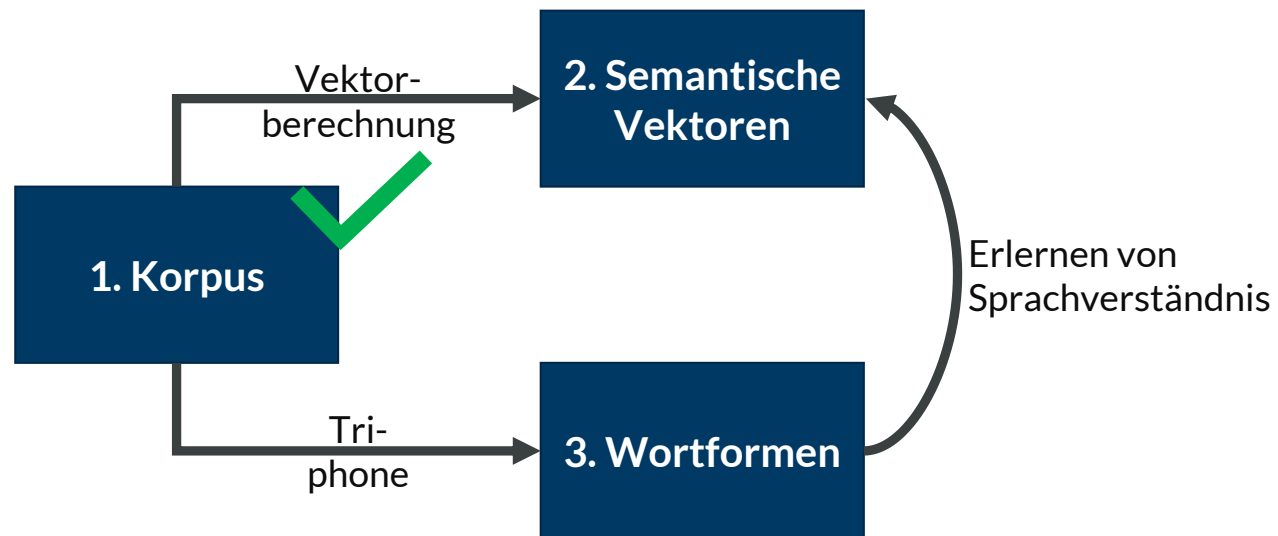
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



# Semantische Vektoren

## Semantische Vektoren

- Anhand des Korpus wurden Vektoren für alle Wortformen sowie deren Flexion via Naive Discriminative Learning (NDL; z. B. Baayen & Ramscar, 2015) berechnet

## Semantische Vektoren

- Anhand des Korpus wurden Vektoren für alle Wortformen sowie deren Flexion via Naive Discriminative Learning (NDL; z. B. Baayen & Ramscar, 2015) berechnet
  - Vektoren auch für Singular/Plural; Generisch/Spezifisch; etc.

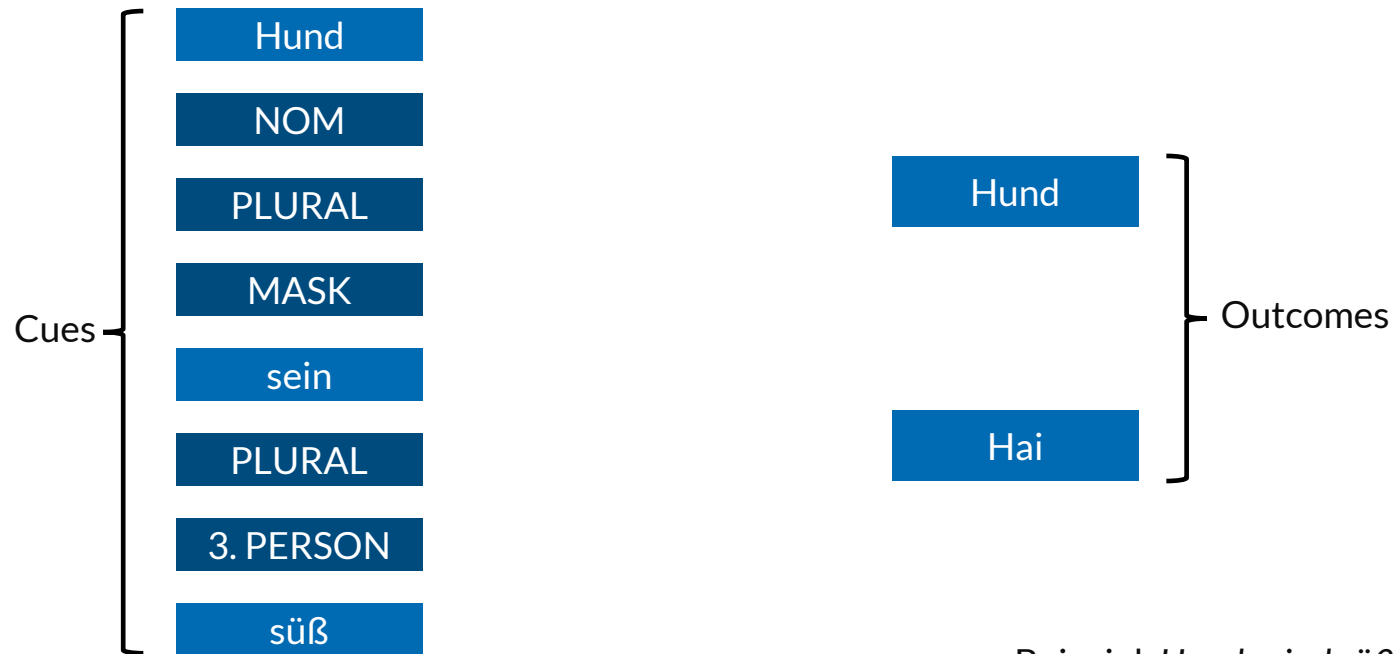
## Semantische Vektoren

- Anhand des Korpus wurden Vektoren für alle Wortformen sowie deren Flexion via Naive Discriminative Learning (NDL; z. B. Baayen & Ramscar, 2015) berechnet
  - Vektoren auch für Singular/Plural; Generisch/Spezifisch; etc.
- NDL folgt dem Rescorla-Wagner-Modell (Rescorla & Wagner, 1972)
  - **Outcomes** (Wortformen /Flexion) werden durch **Cues** (Wortformen/Flexion) vorhergesagt
  - Die Assoziationsstärke zwischen einem Outcome und einem Cue wird stets durch einen Wert dargestellt

## Semantische Vektoren

- Anhand des Korpus wurden Vektoren für alle Wortformen sowie deren Flexion via Naive Discriminative Learning (NDL; z. B. Baayen & Ramscar, 2015) berechnet
  - Vektoren auch für Singular/Plural; Generisch/Spezifisch; etc.
- NDL folgt dem Rescorla-Wagner-Modell (Rescorla & Wagner, 1972)
  - **Outcomes** (Wortformen /Flexion) werden durch **Cues** (Wortformen/Flexion) vorhergesagt
  - Die Assoziationsstärke zwischen einem Outcome und einem Cue wird stets durch einen Wert dargestellt
- Alle **Outcomes** wurden anhand aller **Cues** im selben Satz vorhergesagt

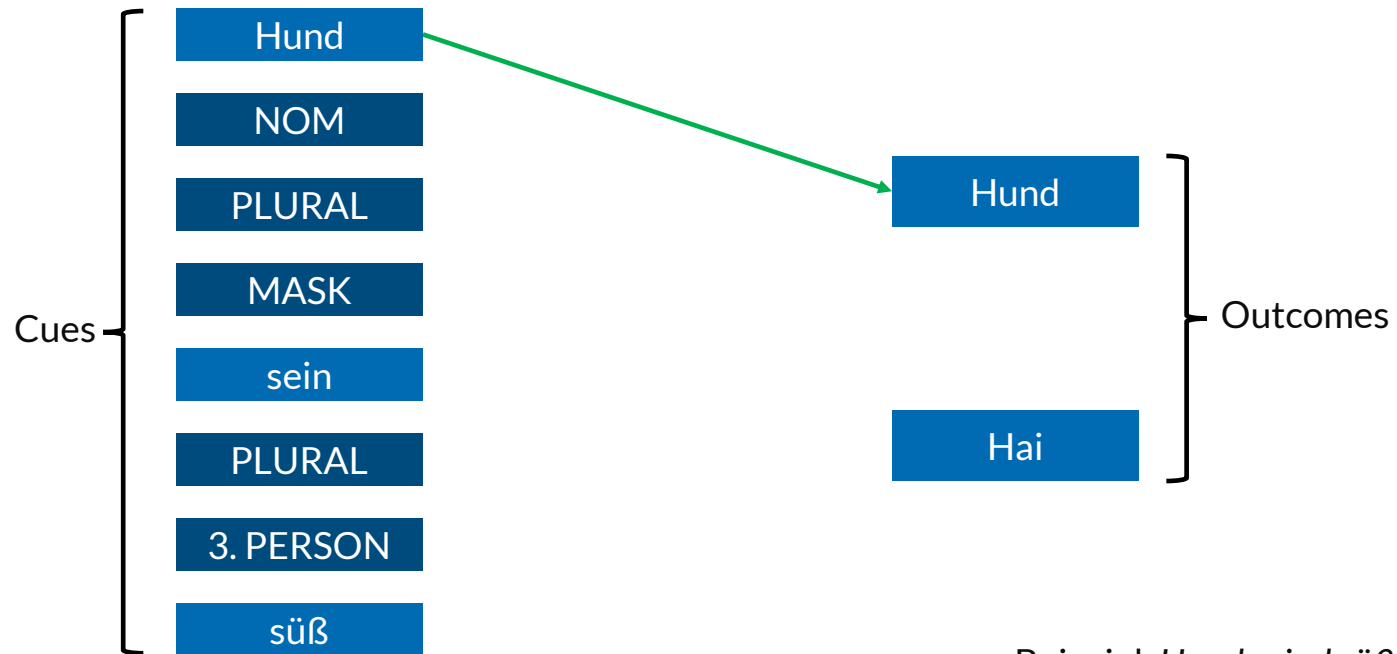
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund							
Hai							

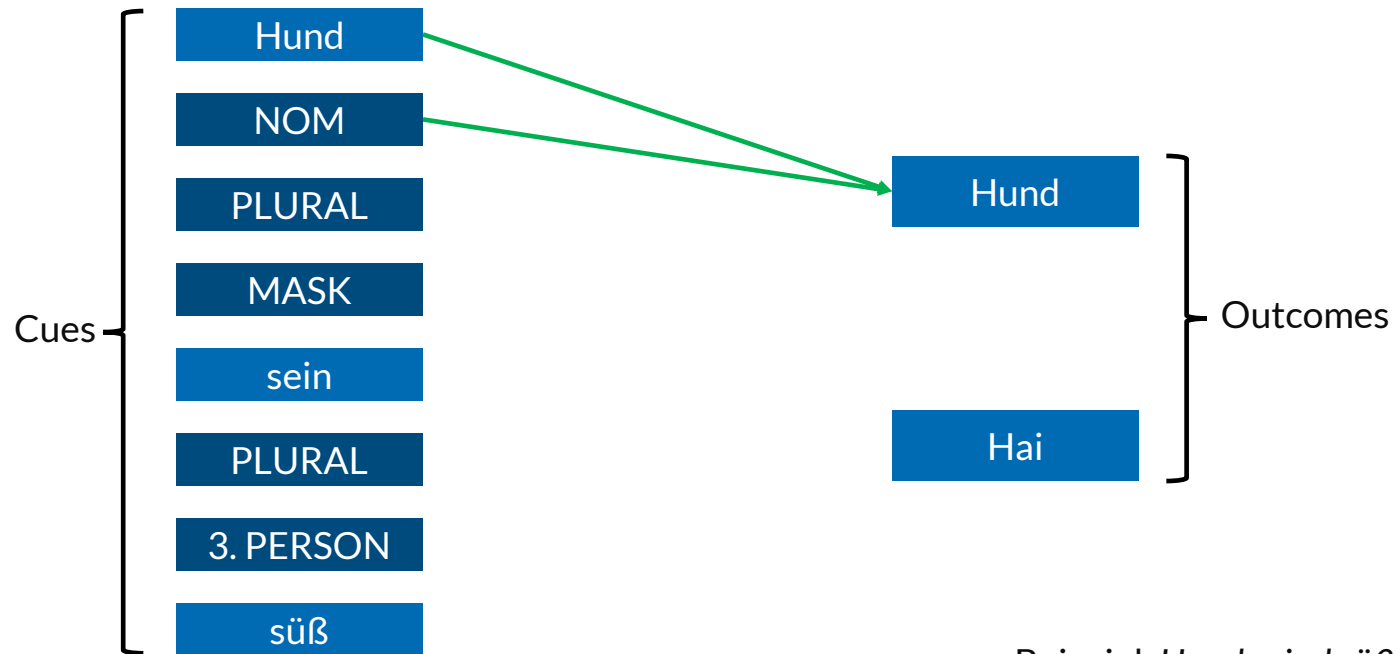
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+						
Hai							

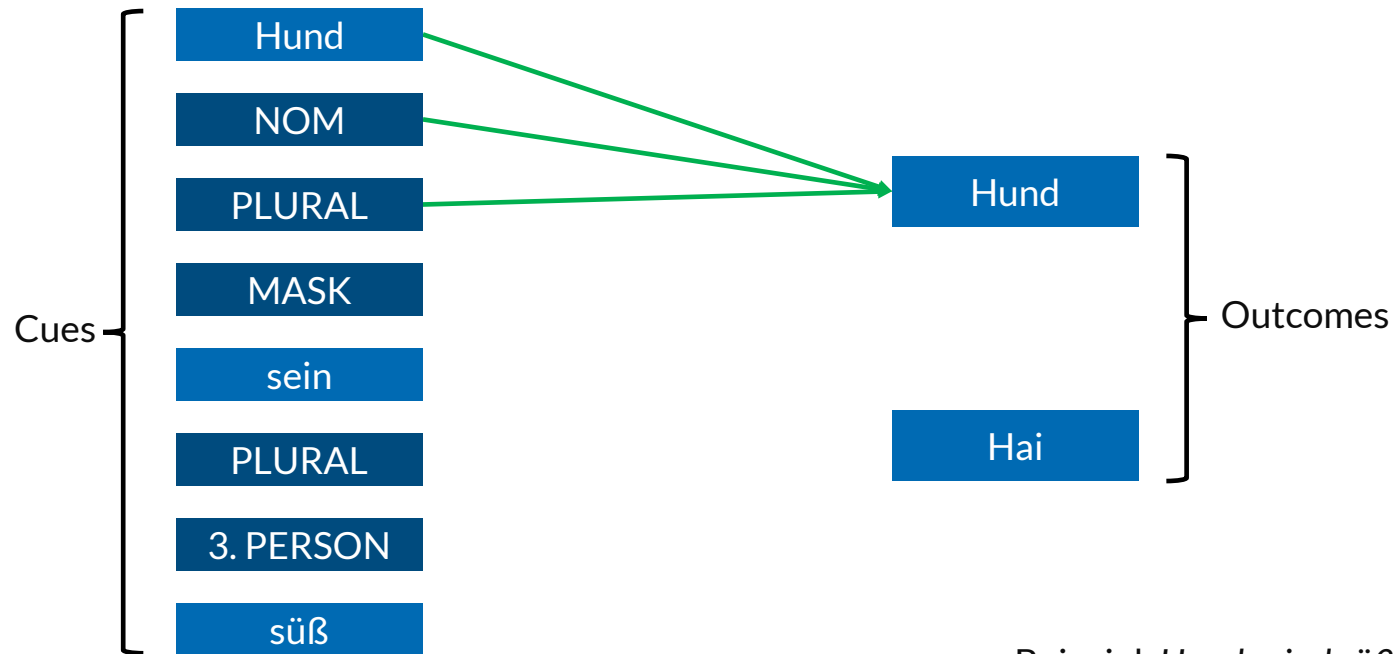
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+					
Hai							

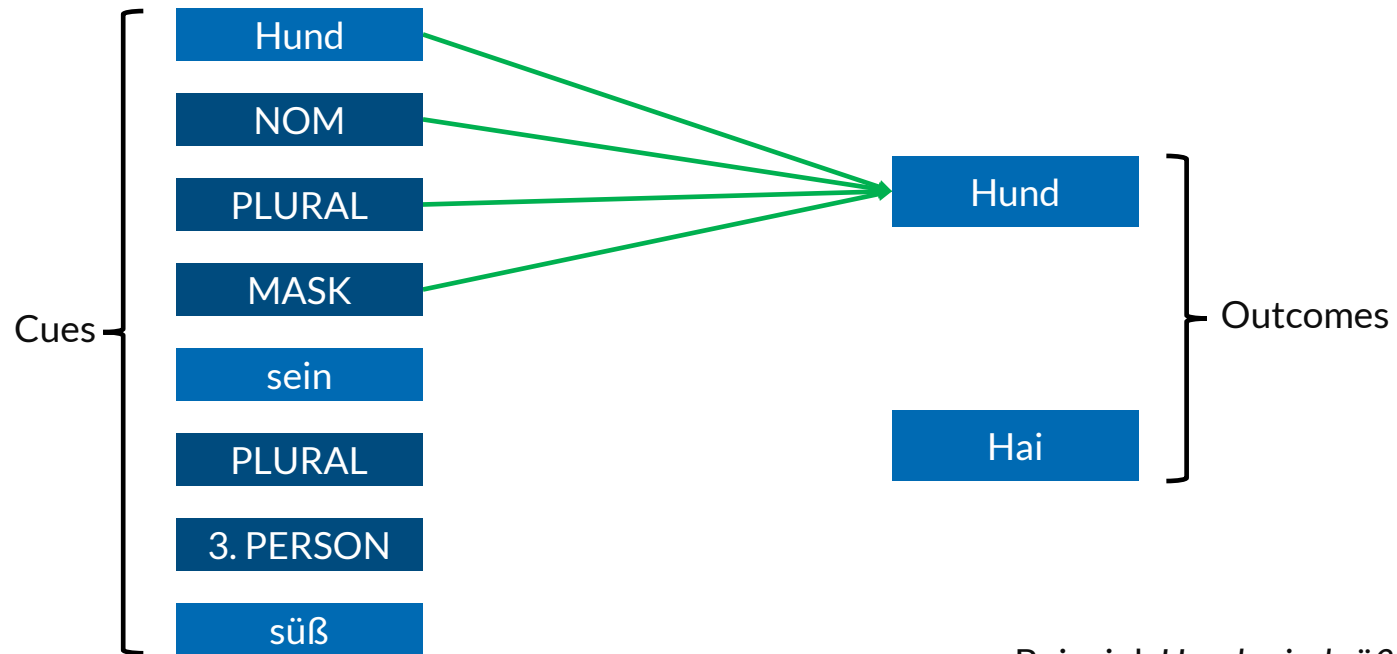
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	+				
Hai							

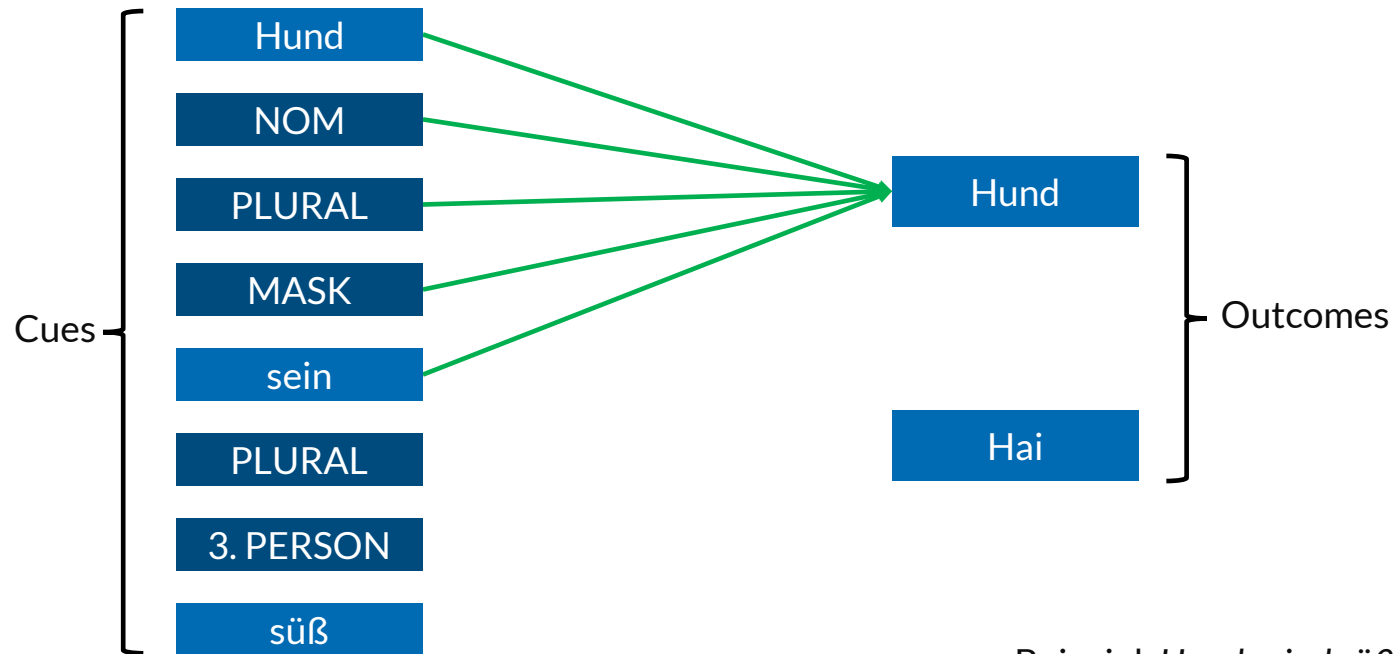
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	+	+			
Hai							

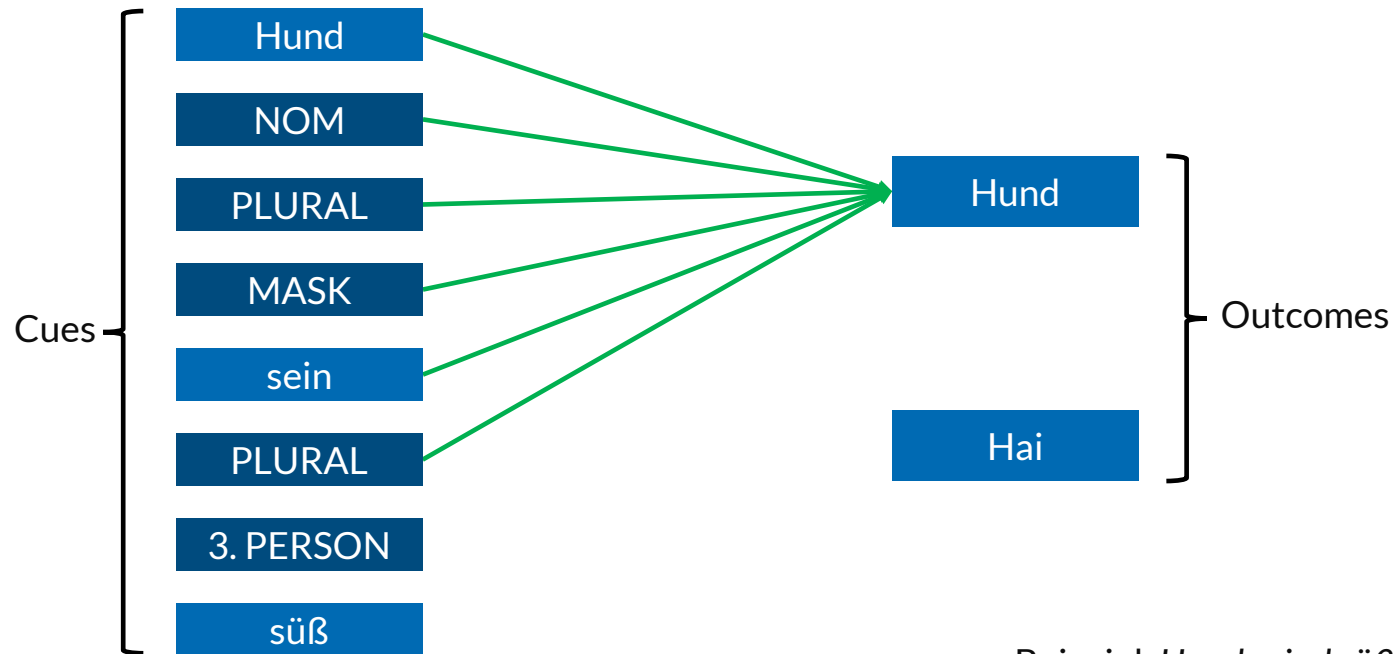
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	+	+	+		
Hai							

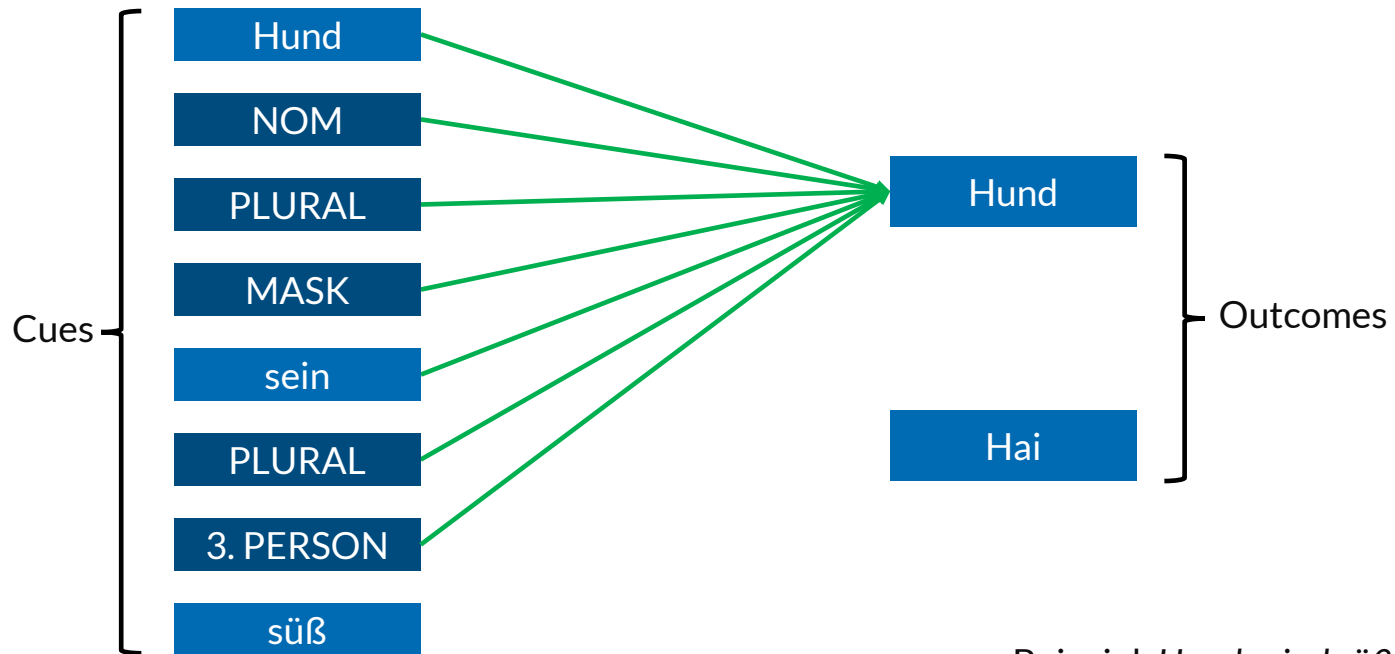
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+		
Hai							

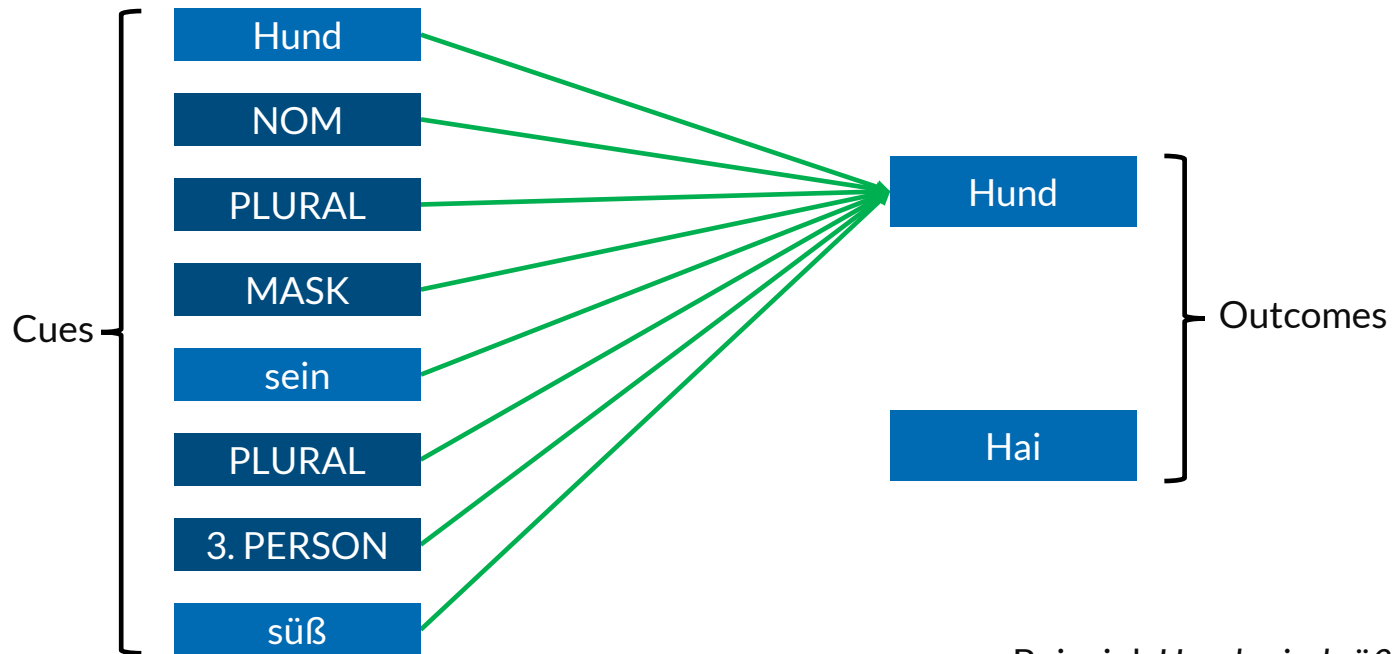
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+	+	
Hai							

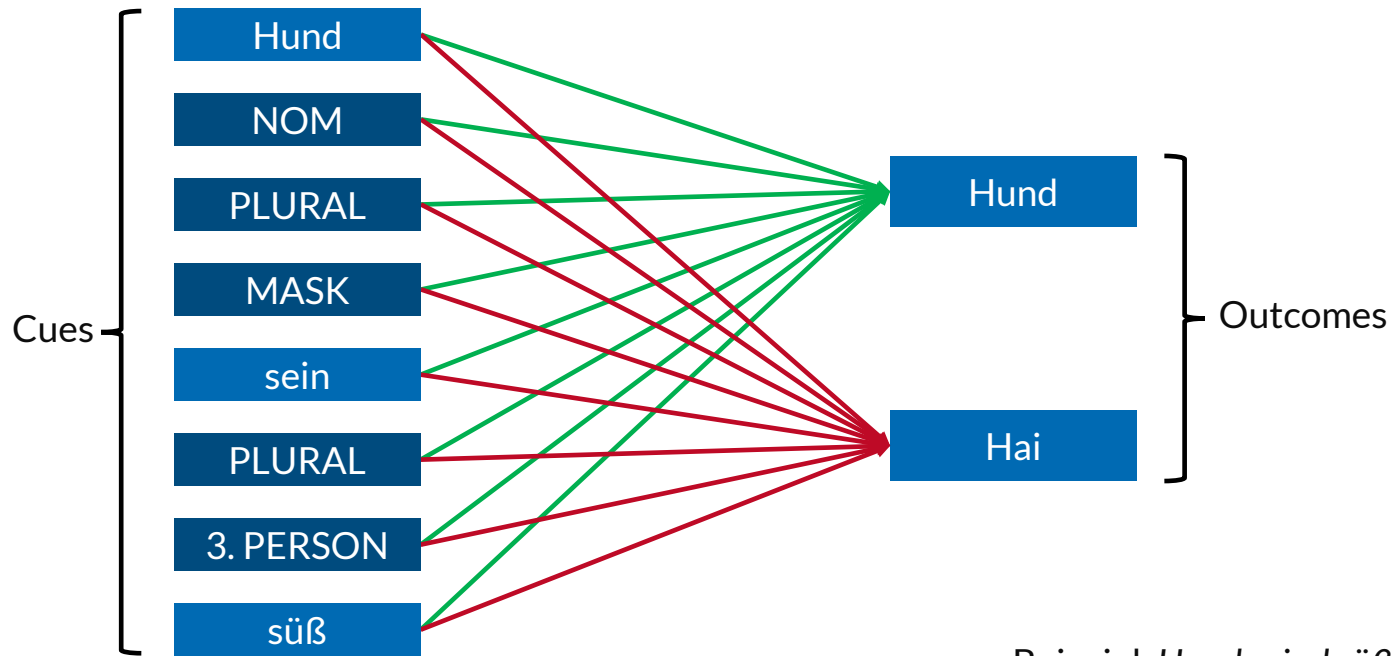
## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+	+	+
Hai							

## NDL – Beispiel



Beispiel: *Hunde sind süß.*

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+	+	+
Hai	-	-	-	-	-	-	-

## NDL → Semantische Vektoren

## NDL → Semantische Vektoren

- Dieser Vorhersageprozess wurde für die 830.000 Sätze des Korpus durchgeführt

## NDL → Semantische Vektoren

- Dieser Vorhersageprozess wurde für die 830.000 Sätze des Korpus durchgeführt
- Als Resultat liegen Assoziationsgewichte vor für
  - Zielwörter
  - Flexion
  - weitere Wörter

## NDL → Semantische Vektoren

- Dieser Vorhersageprozess wurde für die 830.000 Sätze des Korpus durchgeführt
- Als Resultat liegen Assoziationsgewichte vor für
  - Zielwörter
  - Flexion
  - weitere Wörter
- Anhand dieser Reihen von Assoziationsgewichten erhalten wir semantische Vektoren einzelner Wörter und Flexionsfunktionen

## NDL → Semantische Vektoren

- Dieser Vorhersageprozess wurde für die 830.000 Sätze des Korpus durchgeführt
- Als Resultat liegen Assoziationsgewichte vor für
  - Zielwörter
  - Flexion
  - weitere Wörter
- Anhand dieser Reihen von Assoziationsgewichten erhalten wir semantische Vektoren einzelner Wörter und Flexionsfunktionen

	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+	+	+
Hai	-	-	-	-	-	-	-

## NDL → Semantische Vektoren

- Dieser Vorhersageprozess wurde für die 830.000 Sätze des Korpus durchgeführt
- Als Resultat liegen Assoziationsgewichte vor für
  - Zielwörter
  - Flexion
  - weitere Wörter
- Anhand dieser Reihen von Assoziationsgewichten erhalten wir semantische Vektoren einzelner Wörter und Flexionsfunktionen

### Grundbedeutung

<i>Hund</i>	Hund	NOM	PLURAL	MASK	sein	3. PER	süß
Hund	+	+	++	+	+	+	+
Hai	-	-	-	-	-	-	-

## Semantische Vektoren → Rollenbezeichnungen

## Semantische Vektoren → Rollenbezeichnungen

- Anschließend wird die Semantik der Wortformen zusammengesetzt, auch die der untersuchten Rollenbezeichnungen

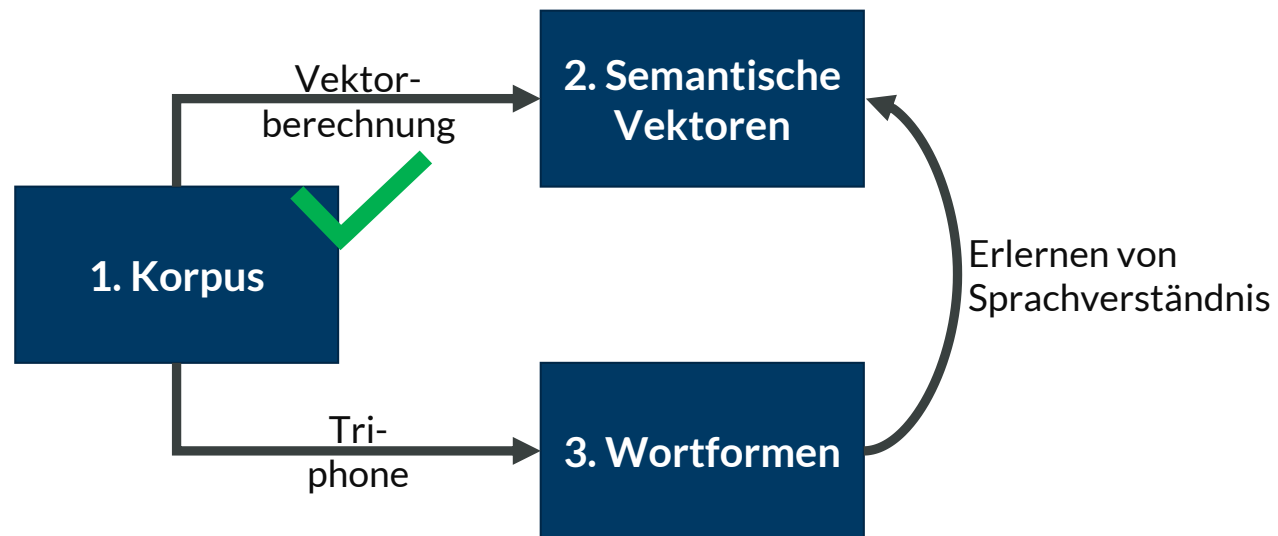
## Semantische Vektoren → Rollenbezeichnungen

- Anschließend wird die Semantik der Wortformen zusammengefasst, auch die der untersuchten Rollenbezeichnungen
- Summe aus Grundbedeutung, Kasus (hier nicht dargestellt), Numerus, Genus & Generizität – generisch vs. spezifisch

Zielform	Basis		Numerus		Genus		Generizität
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Singular}}$	+	$\overrightarrow{\text{Maskulinum}}$	+	$\overrightarrow{\text{generisch}}$
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Singular}}$	+	$\overrightarrow{\text{Maskulinum}}$	+	$\overrightarrow{\text{spezifisch}}$
Lehrerin	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Singular}}$	+	$\overrightarrow{\text{Femininum}}$	+	$\overrightarrow{\text{spezifisch}}$
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Plural}}$	+	$\overrightarrow{\text{Maskulinum}}$	+	$\overrightarrow{\text{generisch}}$
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Plural}}$	+	$\overrightarrow{\text{Maskulinum}}$	+	$\overrightarrow{\text{spezifisch}}$
Lehrerinnen	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{Plural}}$	+	$\overrightarrow{\text{Femininum}}$	+	$\overrightarrow{\text{spezifisch}}$

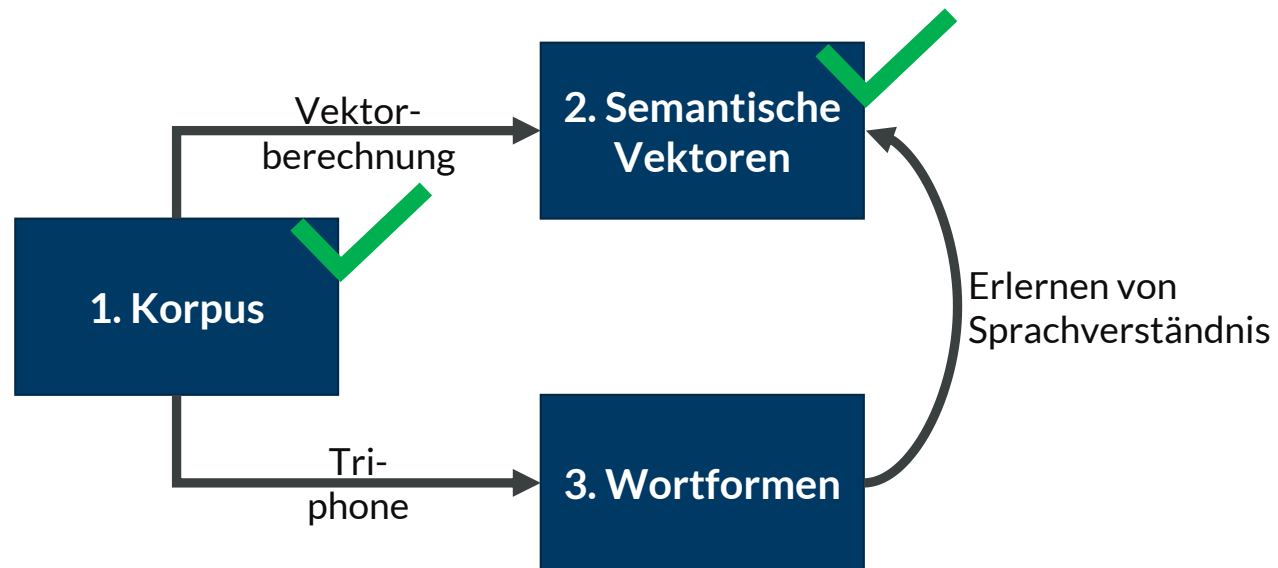
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



## Wortformen

- Wortformen werden in phonologischen Trigrammen (Triphone) repräsentiert
- Trigramme haben in vorherigen Studien zufriedenstellende Resultate gezeigt  
(z. B. Chuang et al., 2020; Schmitz et al., 2021)

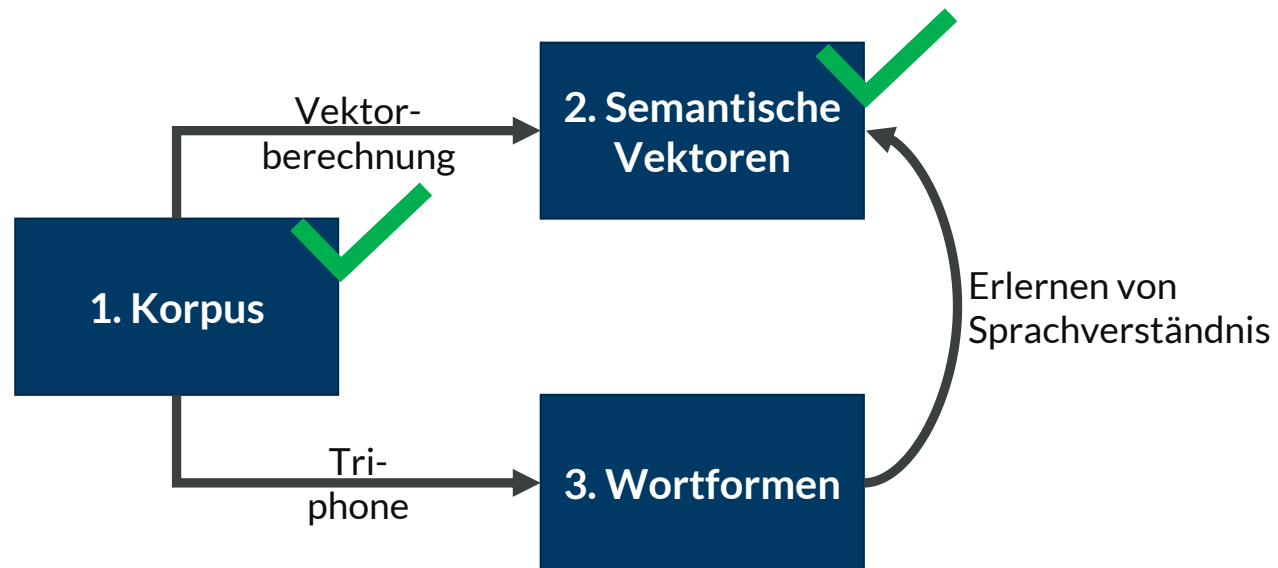
## Wortformen

- Wortformen werden in phonologischen Trigrammen (Triphone) repräsentiert
- Trigramme haben in vorherigen Studien zufriedenstellende Resultate gezeigt  
(z. B. Chuang et al., 2020; Schmitz et al., 2021)

Wortform	#le:	le:r	e:rA	rA#	Arl	rIn	In#
<i>Lehrer</i>	1	1	1	1	0	0	0
<i>Lehrer</i>	1	1	1	1	0	0	0
<i>Lehrerin</i>	1	1	1	0	1	1	1

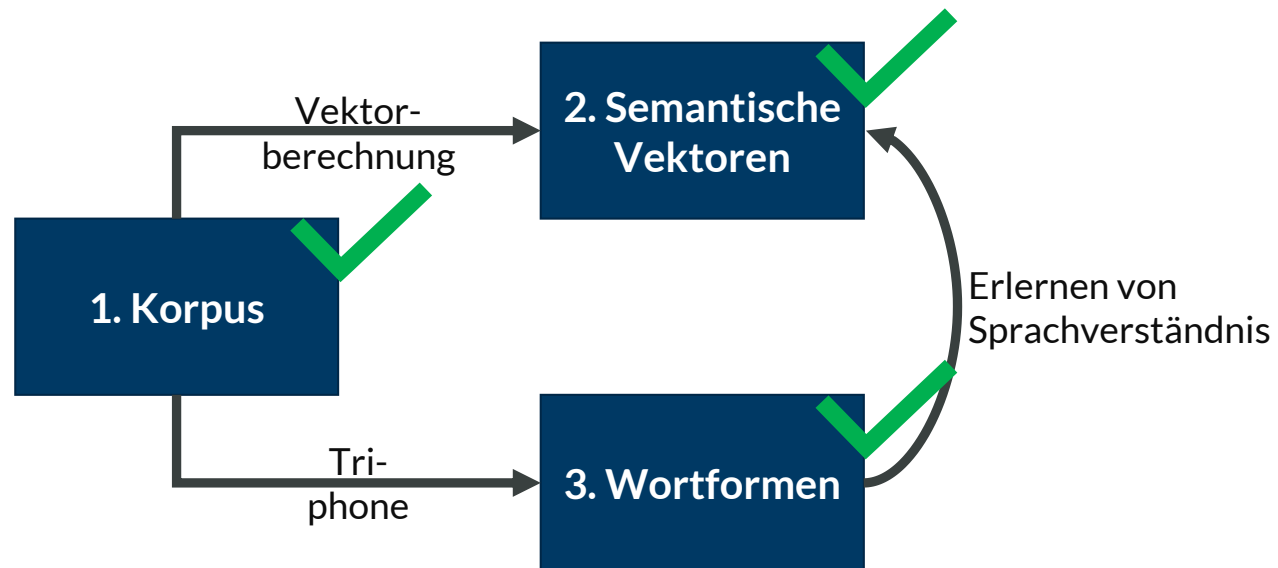
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



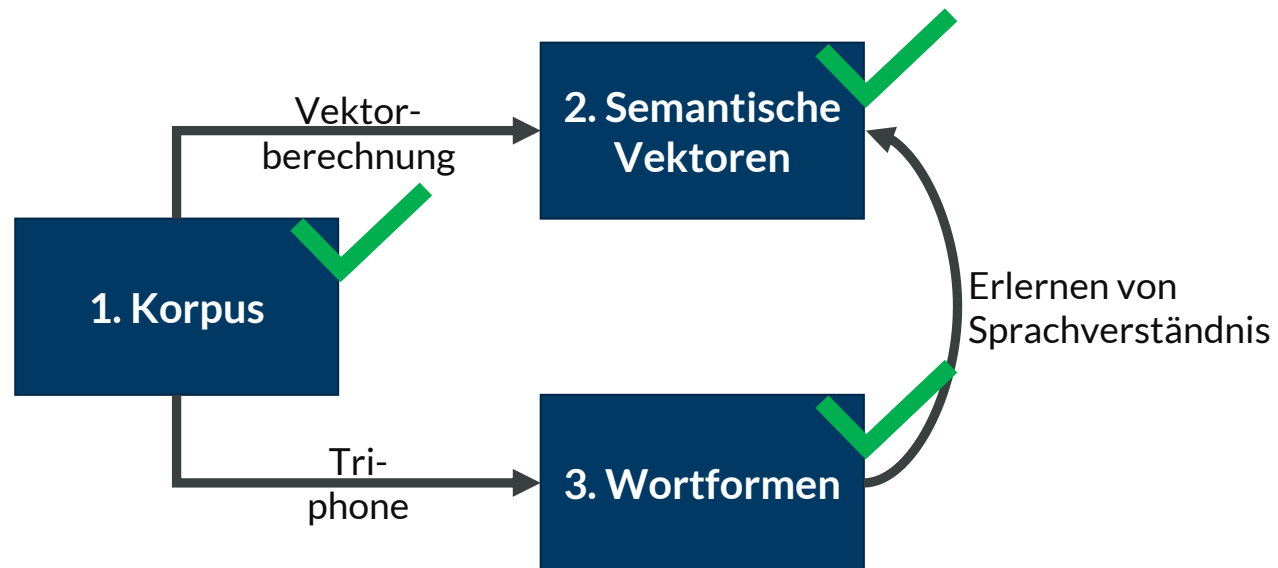
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



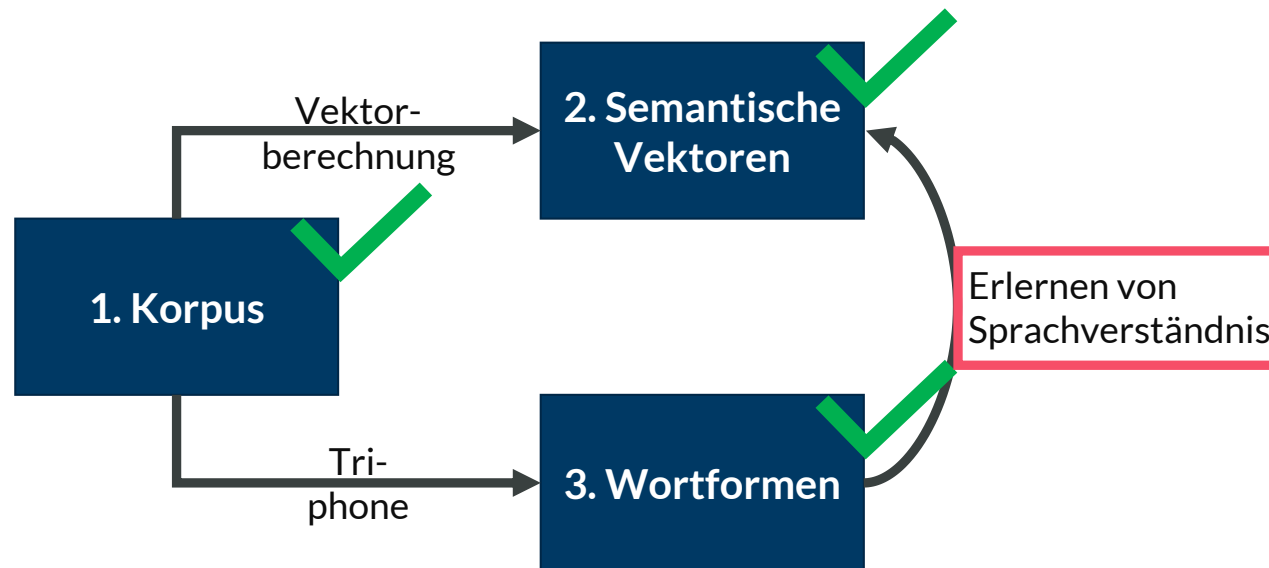
## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (z. B. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt



# Erlernen von Sprachverständnis

## Erlernen von Sprachverständnis

- Das Verständnis wird durch lineares Abbilden der Matrix der Formen auf die Matrix der semantischen Vektoren gelernt

## Erlernen von Sprachverständnis

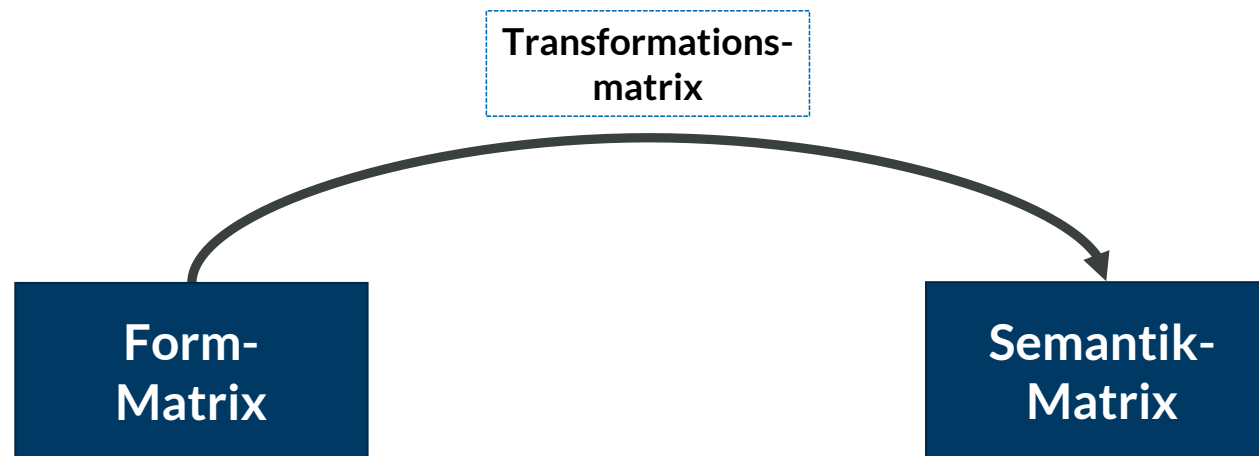
- Das Verständnis wird durch lineares Abbilden der Matrix der Formen auf die Matrix der semantischen Vektoren gelernt

**Form-  
Matrix**

**Semantik-  
Matrix**

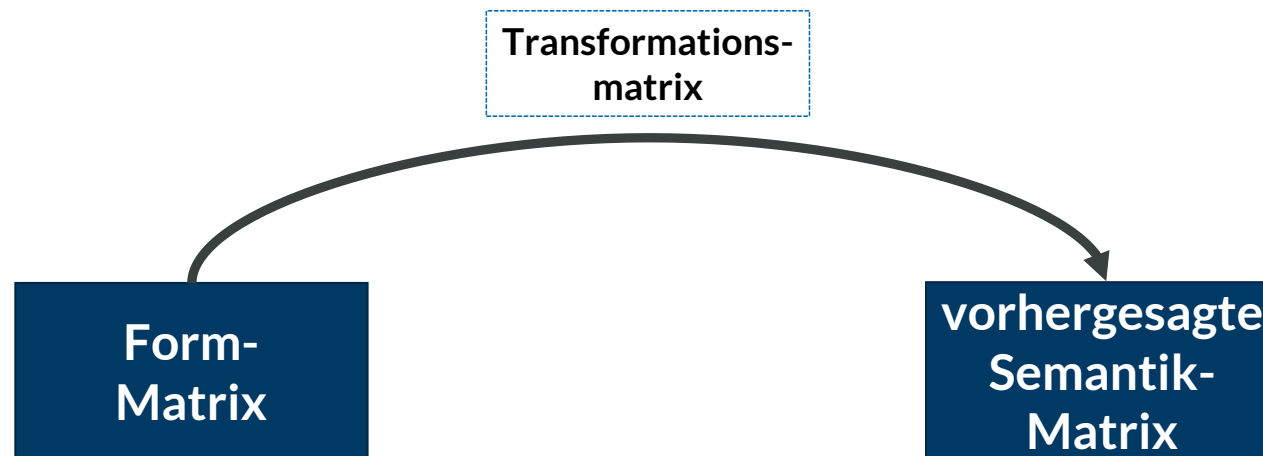
## Erlernen von Sprachverständnis

- Das Verständnis wird durch lineares Abbilden der Matrix der Formen auf die Matrix der semantischen Vektoren gelernt



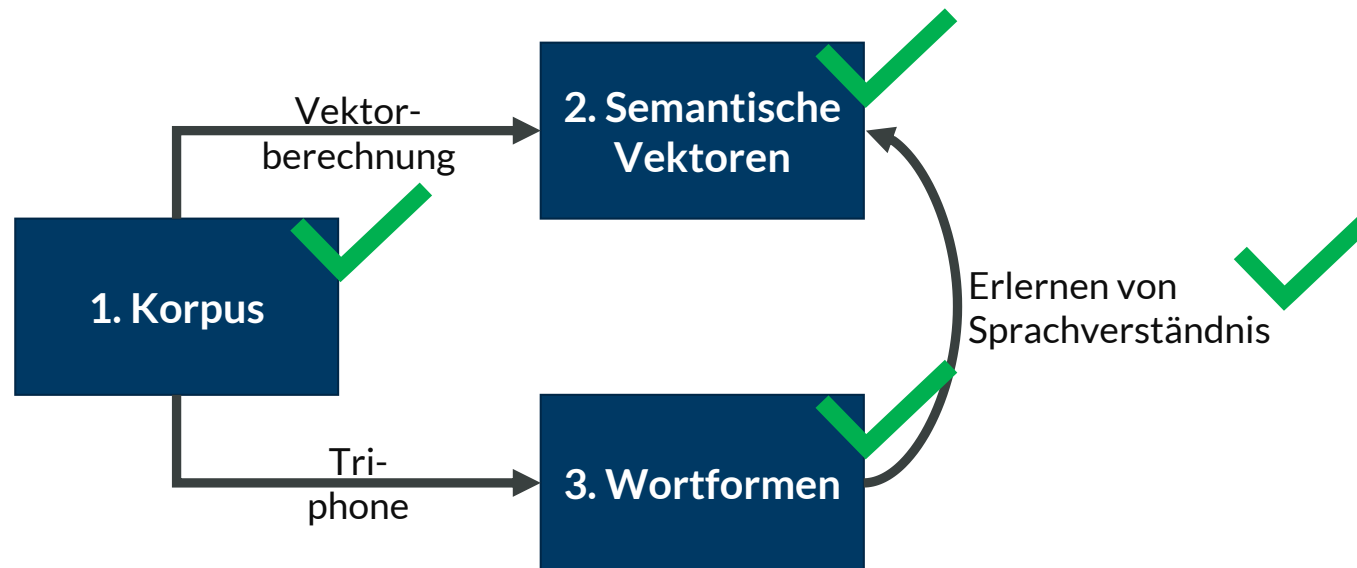
## Erlernen von Sprachverständnis

- Das Verständnis wird durch lineares Abbilden der Matrix der Formen auf die Matrix der semantischen Vektoren gelernt



## Grundidee

- Simulation eines mentalen Lexikons durch die Implementierung eines Linear Discriminative Learning Netzwerks (e.g. Baayen et al., 2019)
- Für die Implementierung werden 3 Dinge benötigt





# Analyse

## Multinominale Logistische Regression

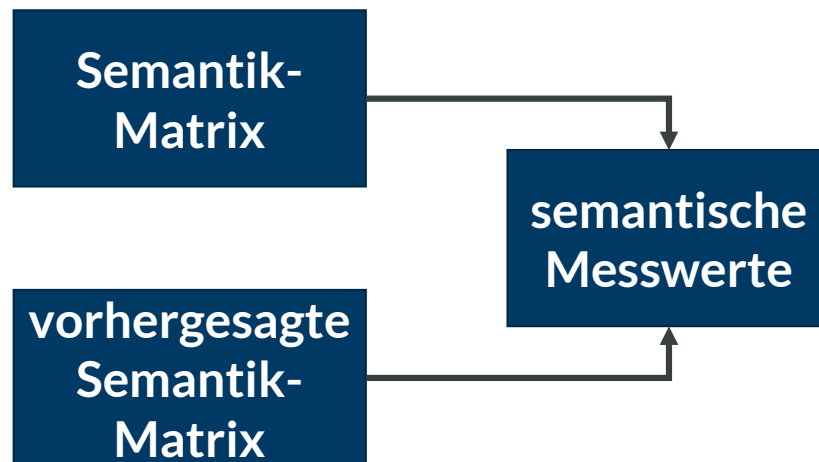
# Semantische Messwerte

## Semantische Messwerte

- Anhand der ursprünglichen und vorhergesagten semantischen Vektoren lassen sich semantische Messwerte berechnen

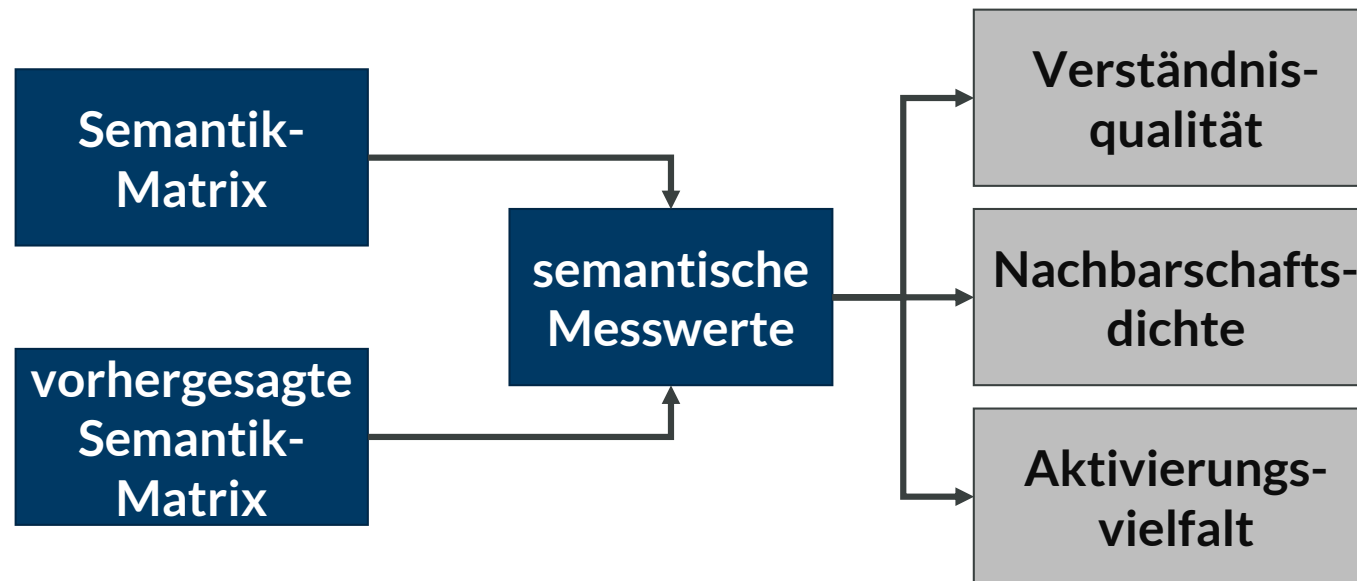
## Semantische Messwerte

- Anhand der ursprünglichen und vorhergesagten semantischen Vektoren lassen sich semantische Messwerte berechnen



## Semantische Messwerte

- Anhand der ursprünglichen und vorhergesagten semantischen Vektoren lassen sich semantische Messwerte berechnen





# Variablen

# Variablen

- semantische Messwerte

# Variablen

- semantische Messwerte
  - **VERSTÄNDNISQUALITÄT**  
Korrelation des ursprünglichen mit dem vorhergesagten Vektor  
höher = besseres Verständnis

# Variablen

- semantische Messwerte
  - **VERSTÄNDNISQUALITÄT**  
Korrelation des ursprünglichen mit dem vorhergesagten Vektor  
höher = besseres Verständnis
  - **NACHBARSCHAFTSDICHTE**  
durchschnittliche Korrelation des Zielwortes mit seinen 8 nächsten Nachbarn  
höher = engere Nachbarschaft

# Variablen

- semantische Messwerte
  - **VERSTÄNDNISQUALITÄT**  
Korrelation des ursprünglichen mit dem vorhergesagten Vektor  
höher = besseres Verständnis
  - **NACHBARSCHAFTSDICHTE**  
durchschnittliche Korrelation des Zielwortes mit seinen 8 nächsten Nachbarn  
höher = engere Nachbarschaft
  - **AKTIVIERUNGSVIELFALT**  
Euklidische Norm des vorhergesagten Vektors  
höher = höheres Maß an Koaktivierung

# Variablen

- semantische Messwerte
  - **VERSTÄNDNISQUALITÄT**  
Korrelation des ursprünglichen mit dem vorhergesagten Vektor  
höher = besseres Verständnis
  - **NACHBARSCHAFTSDICHTE**  
durchschnittliche Korrelation des Zielwortes mit seinen 8 nächsten Nachbarn  
höher = engere Nachbarschaft
  - **AKTIVIERUNGSVIELFALT**  
Euklidische Norm des vorhergesagten Vektors  
höher = höheres Maß an Koaktivierung
- **STEREOTYPIZITÄTSWERTUNGEN** aus Gabriel et al. (2008)

# Multinominale Logistische Regression

# Multinominale Logistische Regression

- abhängige Variable: TYP

Singular generisches Maskulinum, Singular spezifisches Maskulinum, Singular spezifisches Femininum, Plural generisches Maskulinum, Plural spezifisches Maskulinum, Plural spezifisches Femininum

# Multinominale Logistische Regression

- abhängige Variable: TYP  
Singular generisches Maskulinum, Singular spezifisches Maskulinum, Singular spezifisches Femininum, Plural generisches Maskulinum, Plural spezifisches Maskulinum, Plural spezifisches Femininum
- unabhängige Variablen

# Multinominale Logistische Regression

- abhängige Variable: TYP  
Singular generisches Maskulinum, Singular spezifisches Maskulinum, Singular spezifisches Femininum, Plural generisches Maskulinum, Plural spezifisches Maskulinum, Plural spezifisches Femininum
- unabhängige Variablen
  - **AKTIVIERUNGSVIELFALT**

# Multinominale Logistische Regression

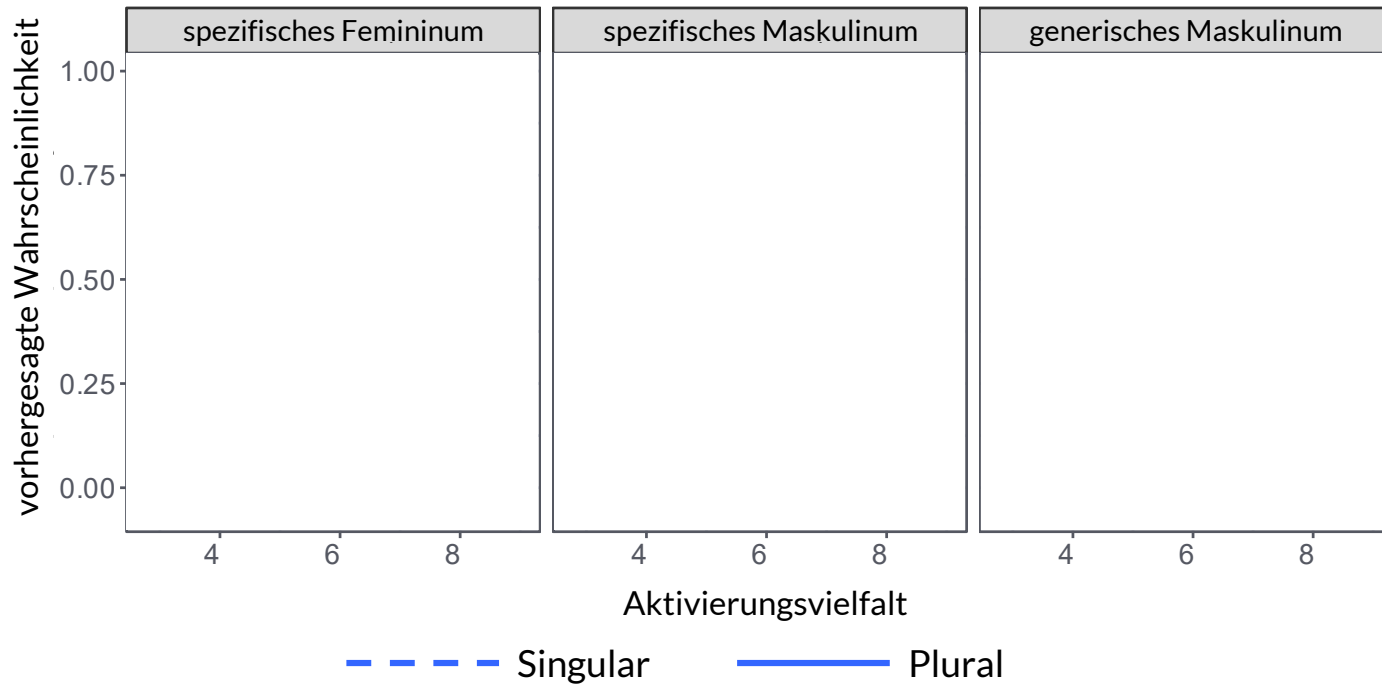
- abhängige Variable: TYP  
Singular generisches Maskulinum, Singular spezifisches Maskulinum, Singular spezifisches Femininum, Plural generisches Maskulinum, Plural spezifisches Maskulinum, Plural spezifisches Femininum
- unabhängige Variablen
  - **AKTIVIERUNGSVIELFALT**
  - **PRINCIPAL COMPONENT (VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE)**

# Multinominale Logistische Regression

- abhängige Variable: TYP  
Singular generisches Maskulinum, Singular spezifisches Maskulinum, Singular spezifisches Femininum, Plural generisches Maskulinum, Plural spezifisches Maskulinum, Plural spezifisches Femininum
- unabhängige Variablen
  - **AKTIVIERUNGSVIELFALT**
  - **PRINCIPAL COMPONENT (VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE)**
  - **STEREOTYPIZITÄTSWERTUNGEN**

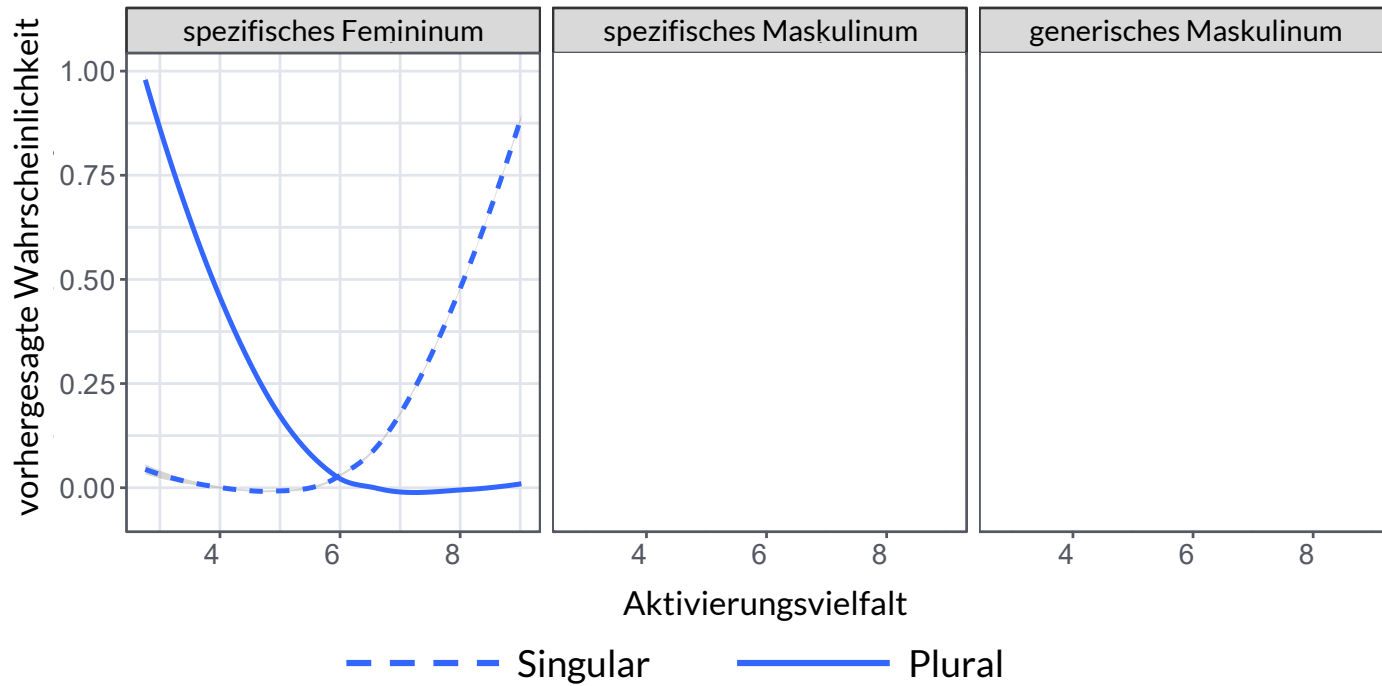
# Ergebnisse

## AKTIVIERUNGSVIELFALT



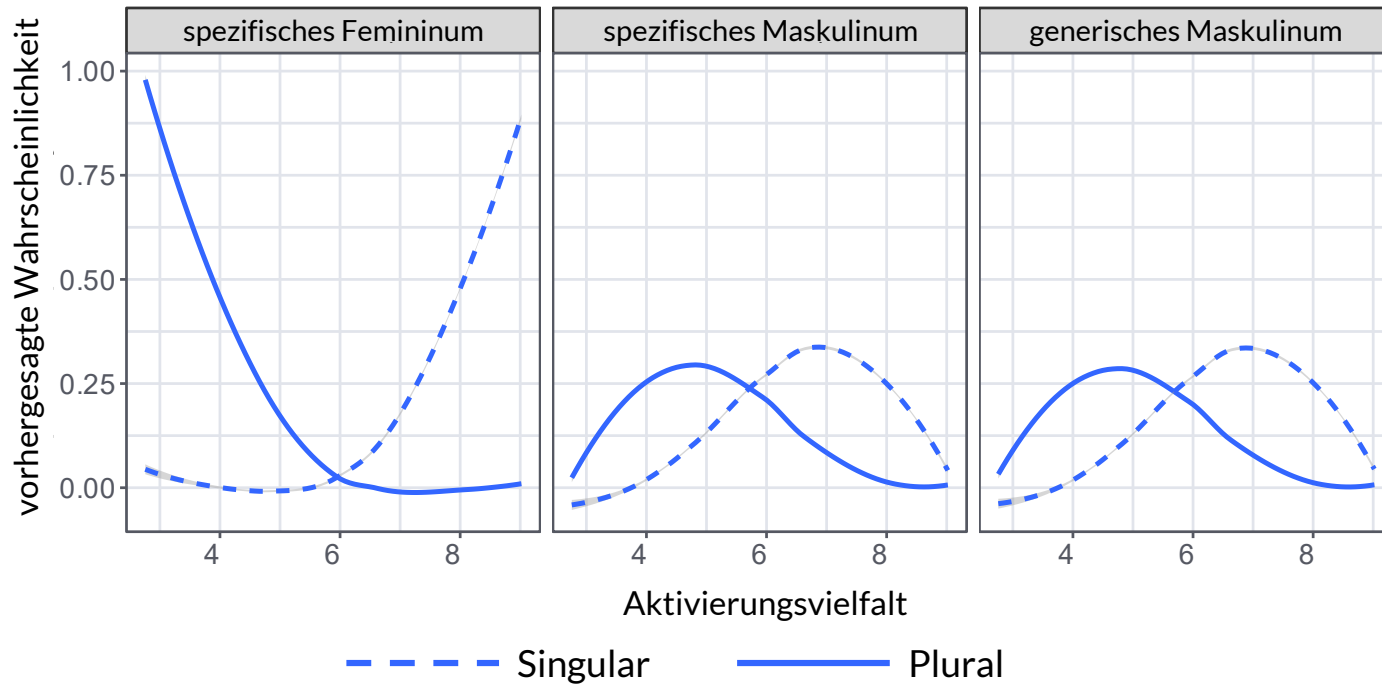
# Ergebnisse

## AKTIVIERUNGSVIELFALT



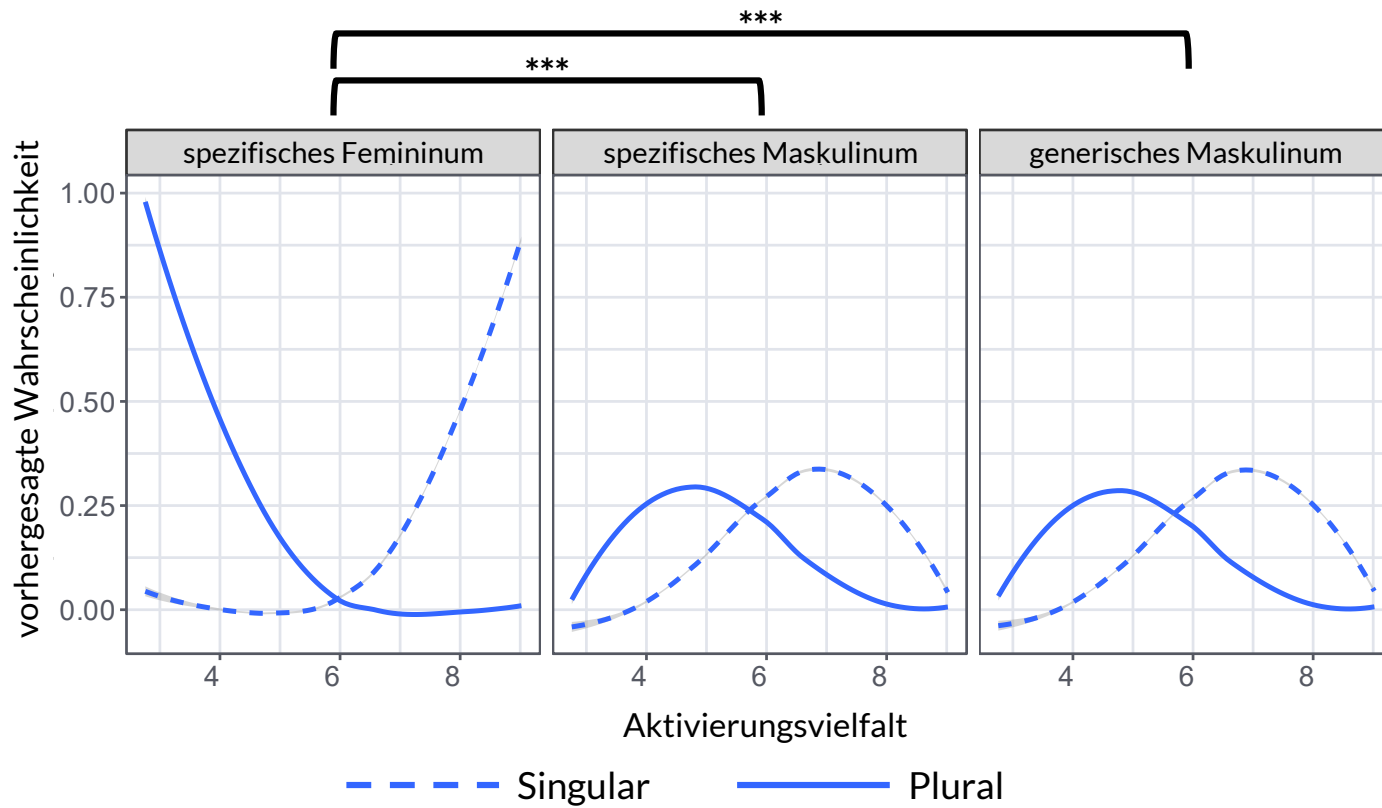
# Ergebnisse

## AKTIVIERUNGSVIELFALT



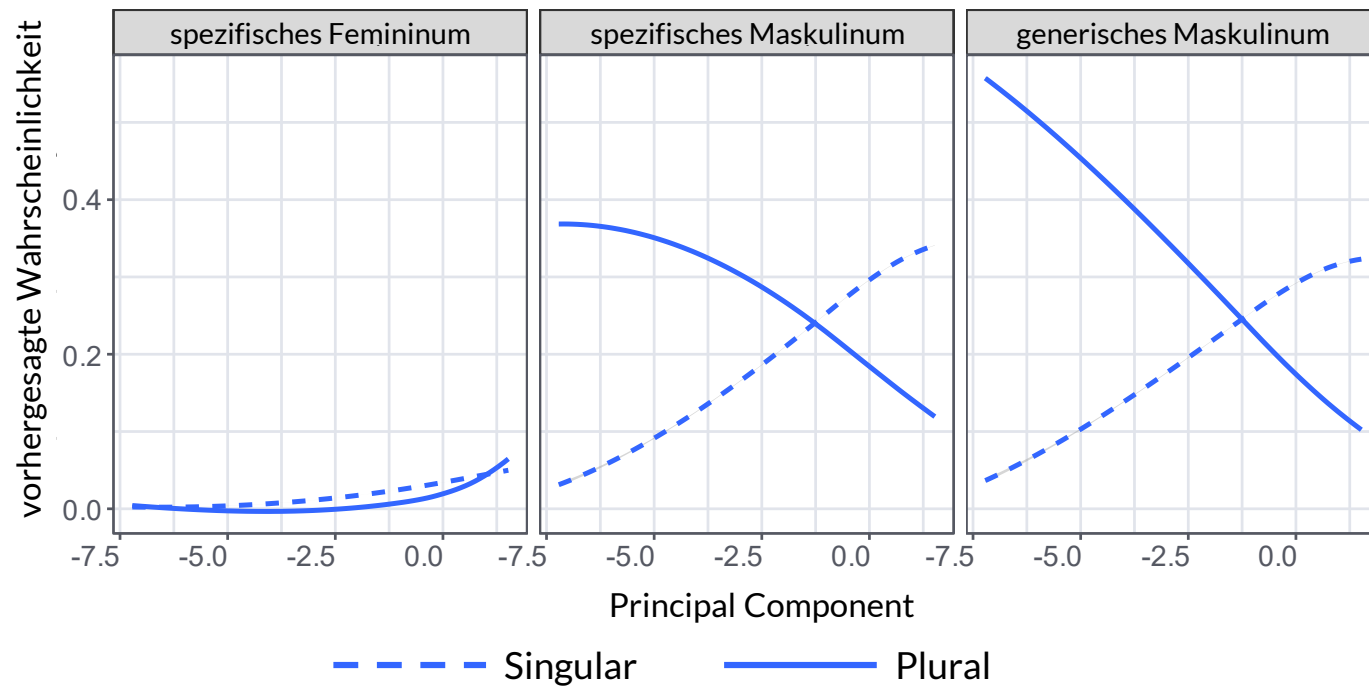
# Ergebnisse

## AKTIVIERUNGSVIELFALT



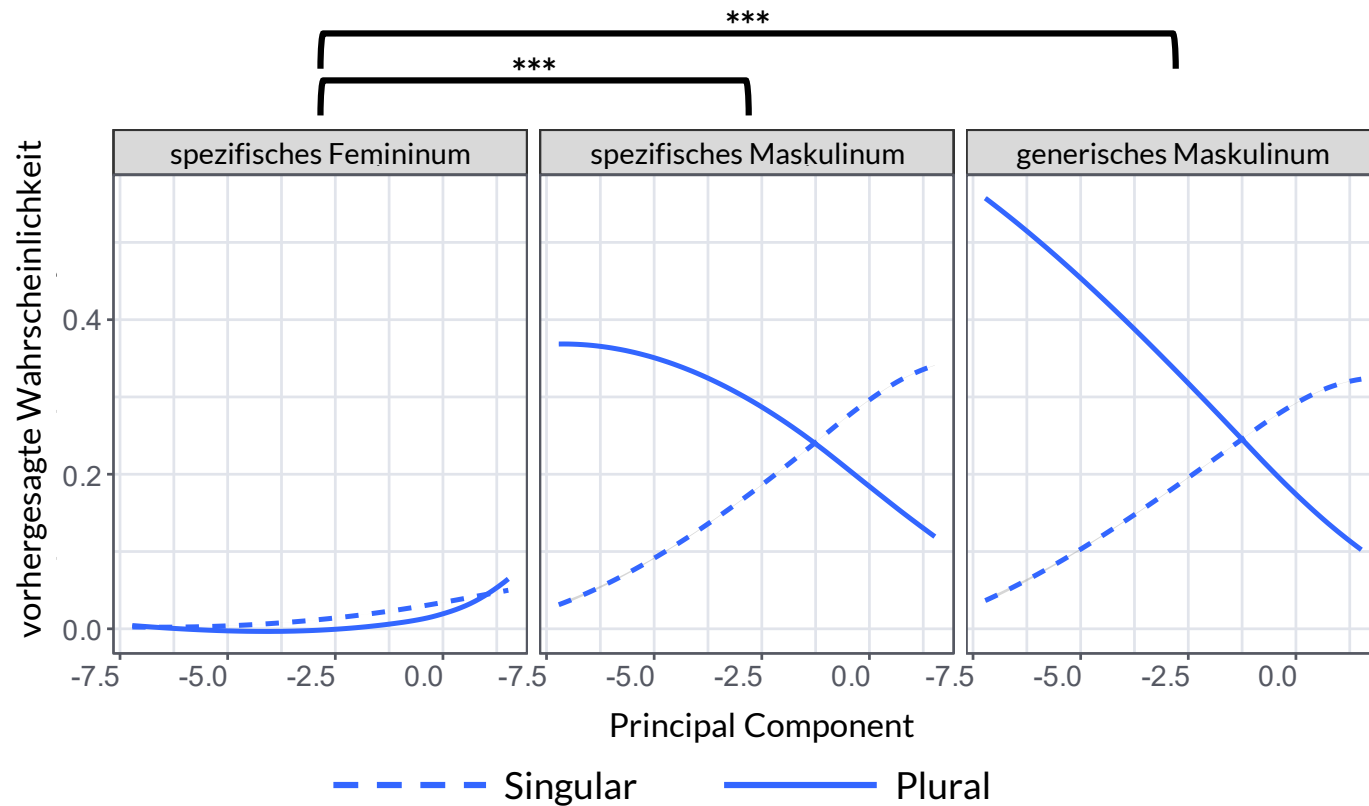
# Ergebnisse

## VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE



# Ergebnisse

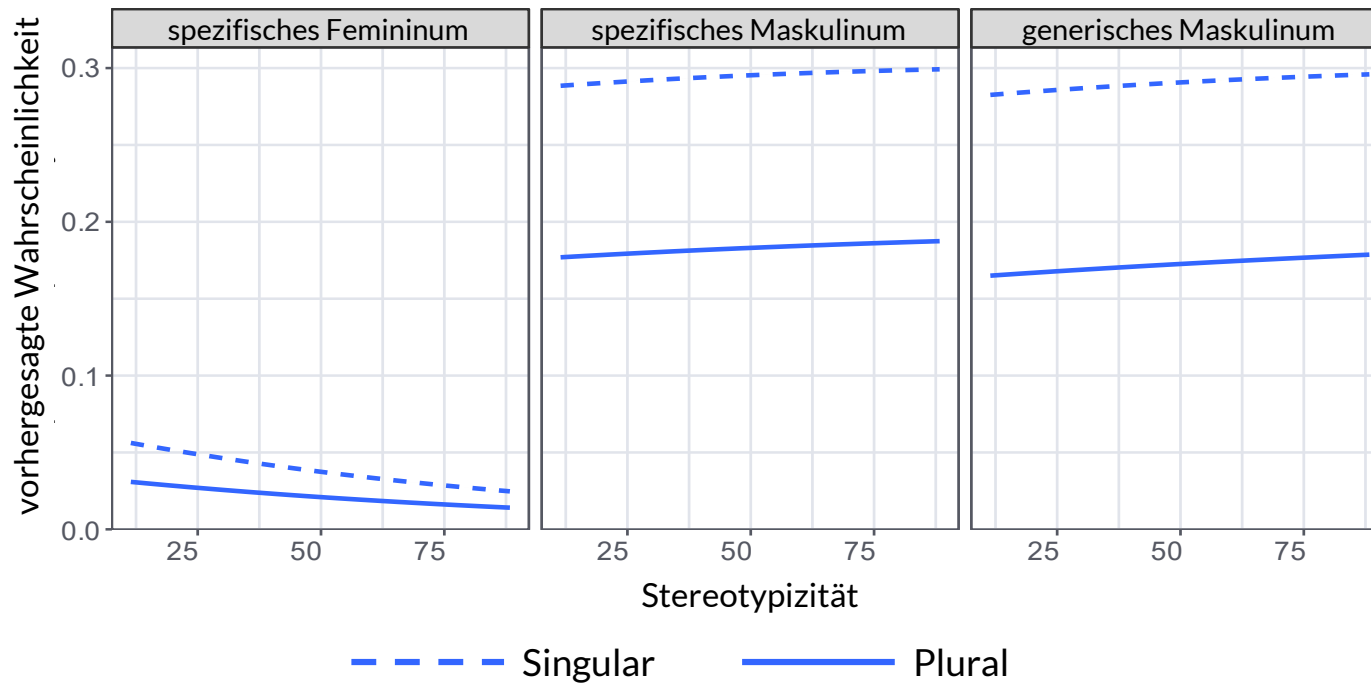
## VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE



# Ergebnisse

## STEREOTYPIZITÄTSWERTUNGEN

kein signifikanter Effekt



# Zusammenfassung

# Zusammenfassung

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

# Zusammenfassung

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

→ ja

# Zusammenfassung

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

→ ja

## Frage 2

Falls ja, inwiefern unterscheidet sich die Semantik generischer Maskulina von der Semantik spezifischer Maskulina und Feminina?

# Zusammenfassung

## Frage 1

Bietet Discriminative Learning einen Einblick in die Semantik generischer Maskulina sowie spezifischer Maskulina und Feminina?

→ ja

## Frage 2

Falls ja, inwiefern unterscheidet sich die Semantik generischer Maskulina von der Semantik spezifischer Maskulina und Feminina?

→ Diskussion



# Diskussion

# Diskussion

- AKTIVIERUNGSVIELFALT

# Diskussion

- **AKTIVIERUNGSVIELFALT**
  - hoch für Feminina im Singular
    - Konkurrenz mit vielen anderen Wörtern, die /In/ enthalten

# Diskussion

- **AKTIVIERUNGSVIELFALT**

- hoch für Feminina im Singular
  - Konkurrenz mit vielen anderen Wörtern, die /In/ enthalten
- medium für Maskulina
  - Konkurrenz mit einigen anderen Wörtern; verschiedene Endungen (-er, -nt, ...)

# Diskussion

- **AKTIVIERUNGSVIELFALT**

- hoch für Feminina im Singular
  - Konkurrenz mit vielen anderen Wörtern, die /ɪn/ enthalten
- medium für Maskulina
  - Konkurrenz mit einigen anderen Wörtern; verschiedene Endungen (-er, -nt, ...)
- niedrig für Feminina im Plural
  - Konkurrenz mit wenigen anderen Wörtern, die /ɪnən/ enthalten

# Diskussion

- **AKTIVIERUNGSVIELFALT**

- hoch für Feminina im Singular
  - Konkurrenz mit vielen anderen Wörtern, die /In/ enthalten
- medium für Maskulina
  - Konkurrenz mit einigen anderen Wörtern; verschiedene Endungen (-er, -nt, ...)
- niedrig für Feminina im Plural
  - Konkurrenz mit wenigen anderen Wörtern, die /Inən/ enthalten

- **VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE**

# Diskussion

- **AKTIVIERUNGSVIELFALT**

- hoch für Feminina im Singular
  - Konkurrenz mit vielen anderen Wörtern, die /ɪn/ enthalten
- medium für Maskulina
  - Konkurrenz mit einigen anderen Wörtern; verschiedene Endungen (-er, -nt, ...)
- niedrig für Feminina im Plural
  - Konkurrenz mit wenigen anderen Wörtern, die /ɪnən/ enthalten

- **VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE**

- Feminina „leben“ in ihrem eigenen Teil des semantischen Raums, sind deshalb signifikant anders als Maskulina

# Diskussion

- **AKTIVIERUNGSVIELFALT**

- hoch für Feminina im Singular
  - Konkurrenz mit vielen anderen Wörtern, die /In/ enthalten
- medium für Maskulina
  - Konkurrenz mit einigen anderen Wörtern; verschiedene Endungen (-er, -nt, ...)
- niedrig für Feminina im Plural
  - Konkurrenz mit wenigen anderen Wörtern, die /Inən/ enthalten

- **VERSTÄNDNISQUALITÄT + NACHBARSCHAFTSDICHTE**

- Feminina „leben“ in ihrem eigenen Teil des semantischen Raums, sind deshalb signifikant anders als Maskulina
- alle nächsten Nachbarn der untersuchten Feminina sind untersuchte Feminina



# Fazit

## Fazit

- Der männliche Bias ist auf die ähnlichen semantischen Merkmale von generischen und spezifischen Maskulina zurückzuführen

## Fazit

- Der männliche Bias ist auf die ähnlichen semantischen Merkmale von generischen und spezifischen Maskulina zurückzuführen
- Ursprung des männlichen Bias ist also das Sprachsystem selbst, kein äußerer Einfluss

## Fazit

- Der männliche Bias ist auf die ähnlichen semantischen Merkmale von generischen und spezifischen Maskulina zurückzuführen
- Ursprung des männlichen Bias ist also das Sprachsystem selbst, kein äußerer Einfluss
- Die präsentierten Ergebnisse stimmen mit denen vorheriger Untersuchungen überein (z. B. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)

## Fazit

- Der männliche Bias ist auf die ähnlichen semantischen Merkmale von generischen und spezifischen Maskulina zurückzuführen
- Ursprung des männlichen Bias ist also das Sprachsystem selbst, kein äußerer Einfluss
- Die präsentierten Ergebnisse stimmen mit denen vorheriger Untersuchungen überein (z. B. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001)
- Zukünftige Forschung wird zeigen
  - ob die für unsere Daten berechneten LDL-Maße eine Vorhersagekraft für Verhaltensmaße haben, z.B. Reaktionszeiten
  - wie sich (neue) neutralere Formen, z.B. *Lehrer\*innen*, *LehrerInnen*, verhalten



**Vielen Dank für die Aufmerksamkeit!**

weitere Details unter [psyarxiv.com/c27r9](https://psyarxiv.com/c27r9)

## Referenzen 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, 1–39. <https://doi.org/10.1155/2019/4895891>
- Baayen, R. H., & Ramscar, M. (2015). Abstraction, storage and naive discriminative learning. *Handbook of Cognitive Linguistics*, 39, 100–120. <https://doi.org/10.1515/9783110292022-006>
- 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>
- Demarmels, S. (2017). „Gesucht: Assistentin oder Sekretär der Geschäftsleitung“ – Gendersensitive Formulierungen in Stellenanzeigen aus der Perspektive der Textsorte. In *Stellenanzeigen als Instrument des Employer Branding in Europa*. [https://doi.org/10.1007/978-3-658-12719-0\\_11](https://doi.org/10.1007/978-3-658-12719-0_11)
- Gabriel, U., Gygas, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, 40(1), 206–212. <https://doi.org/10.3758/BRM.40.1.206>
- Garnham, A., Gabriel, U., Sarrasin, O., Gygas, P., & Oakhill, J. (2012). Gender Representation in Different Languages and Grammatical Marking on Pronouns: When Beauticians, Musicians, and Mechanics Remain Men. *Discourse Processes*, 49(6), 481–500. <https://doi.org/10.1080/0163853X.2012.688184>
- Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*.
- Gygas, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23(3), 464–485. <https://doi.org/10.1080/01690960701702035>
- Gygas, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: a challenge for our cognitive system. *Language Sciences*, 83, 101328. <https://doi.org/10.1016/j.langsci.2020.101328>

## Referenzen 2/2

- Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: a challenge for our cognitive system. *Language Sciences*, 83, 101328. <https://doi.org/10.1016/j.langsci.2020.101328>
- Irmen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, 57(5), 367–375. <https://doi.org/10.1027/1618-3169/a000044>
- Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. *Zeitschrift Für Psychologie / Journal of Psychology*, 213(3), 167–175. <https://doi.org/10.1026/0044-3409.213.3.167>
- Koch, M. (2021). *Kognitive Effekte des generischen Maskulinums und genderneutraler Alternativen im Deutschen – eine empirische Untersuchung*. Master's Thesis. Technische Universität Braunschweig.
- Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical Gender in German Influences How Role-Nouns Are Interpreted: Evidence from ERPs. *Discourse Processes*, 56(8), 643–654. <https://doi.org/10.1080/0163853X.2018.1541382>
- Schmid, H. (1999). Improvements in part-of-speech tagging with an application to German. In S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, & D. Yarowsky (Eds.), *Natural language processing using very large corpora* (pp. 13–25). Springer. [https://doi.org/10.1007/978-94-017-2390-9\\_2](https://doi.org/10.1007/978-94-017-2390-9_2)
- 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>
- Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, 52(3), 131–140. <https://doi.org/10.1026//0033-3042.52.3.131>
- Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name Your Favorite Musician. *Journal of Language and Social Psychology*, 20(4), 464–469. <https://doi.org/10.1177/0261927X01020004004>

## NDL → Semantische Vektoren

## NDL → Semantische Vektoren

- So wie wir einen Vektor für die Grundbedeutung von *Hund* erhalten, so erhalten wir auch einen Vektor für z. B. Plural

## NDL → Semantische Vektoren

- So wie wir einen Vektor für die Grundbedeutung von *Hund* erhalten, so erhalten wir auch einen Vektor für z. B. Plural
- Um die semantische Repräsentation für *Hunde* zu erhalten, summieren wir die Vektoren von *Hund* und Plural

## NDL → Semantische Vektoren

- *Hund HHuunndd Hund + Plural PPluurraall Plural*
- So wie wir einen Vektor für die Grundbedeutung von *Hund* erhalten, so erhalten wir auch einen Vektor für z. B. Plural
- Um die semantische Repräsentation für *Hunde* zu erhalten, summieren wir die Vektoren von *Hund* und Plural

$$Hunde = unde = \overrightarrow{Hund} + \overrightarrow{Plural}$$

## NDL → Semantische Vektoren

- *Hund HHuunndd Hund + Plural PPluurraall Plural*
- So wie wir einen Vektor für die Grundbedeutung von *Hund* erhalten, so erhalten wir auch einen Vektor für z. B. Plural
- Um die semantische Repräsentation für *Hunde* zu erhalten, summieren wir die Vektoren von *Hund* und Plural

$$Hunde = unde = \overrightarrow{Hund} + \overrightarrow{Plural}$$

- Da wir uns im Deutschen befinden, wird die komplette Summe noch durch Kasus, Numerus und Genus ergänzt, z. B. für *Hundes*

## NDL → Semantische Vektoren

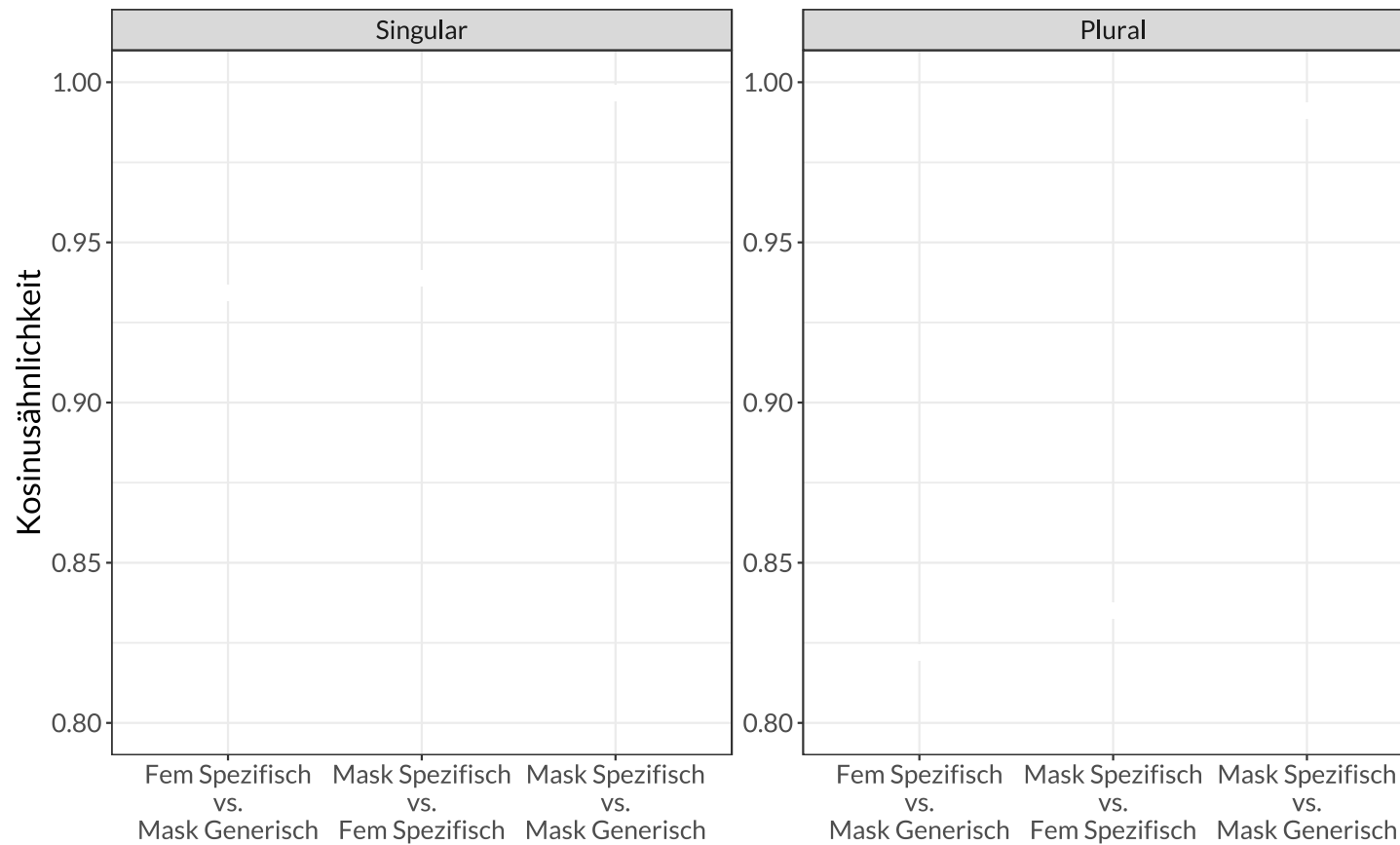
- *Hund HHuunndd Hund + Genitiv GGeenniittiivv Genitiv + Singular  
SSiinngguullaarr Singular + Maskulinum MMAasskkuulliinnuumm Maskulinum*
- *Hund HHuunndd Hund + Plural PPluurraall Plural*
- So wie wir einen Vektor für die Grundbedeutung von *Hund* erhalten, so erhalten wir auch einen Vektor für z. B. Plural
- Um die semantische Repräsentation für *Hunde* zu erhalten, summieren wir die Vektoren von *Hund* und Plural

$$\overrightarrow{Hunde} = \overrightarrow{unde} = \overrightarrow{Hund} + \overrightarrow{Plural}$$

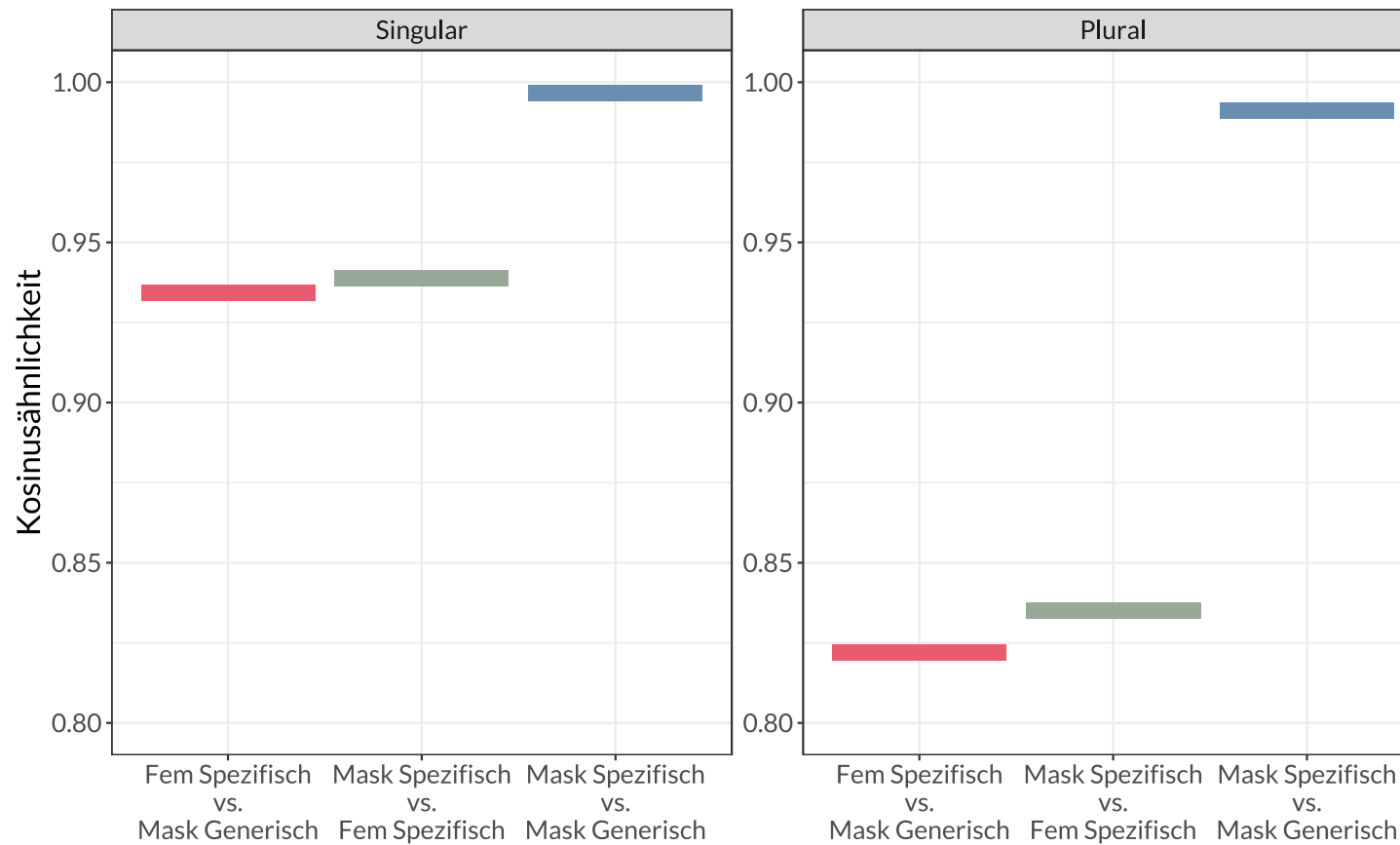
- Da wir uns im Deutschen befinden, wird die komplette Summe noch durch Kasus, Numerus und Genus ergänzt, z. B. für *Hundes*

# Semantische Ähnlichkeit

# Semantische Ähnlichkeit



# Semantische Ähnlichkeit





# Diskussion

## Diskussion

- Unsere Ergebnisse stehen im Einklang mit den Annahmen früherer Untersuchungen

## Diskussion

- Unsere Ergebnisse stehen im Einklang mit den Annahmen früherer Untersuchungen
  - Stahlberg et al. (2001)  
das maskuline Geschlecht von [maskulinen] Generika hat eine **semantische Komponente der "Männlichkeit"**

## Diskussion

- Unsere Ergebnisse stehen im Einklang mit den Annahmen früherer Untersuchungen
  - Stahlberg et al. (2001)  
das maskuline Geschlecht von [maskulinen] Generika hat eine **semantische Komponente der "Männlichkeit"**
  - Irmen & Linner (2005)  
semantische Ähnlichkeit von generischen und spezifischen Maskulina aufgrund ihrer **Resonanz mit dem Lexikon und untereinander**

## Diskussion

- Unsere Ergebnisse stehen im Einklang mit den Annahmen früherer Untersuchungen
  - Stahlberg et al. (2001)  
das maskuline Geschlecht von [maskulinen] Generika hat eine **semantische Komponente der "Männlichkeit"**
  - Irmen & Linner (2005)  
semantische Ähnlichkeit von generischen und spezifischen Maskulina aufgrund ihrer **Resonanz mit dem Lexikon und untereinander**
  - Gygax et al. (2012) und Gygax et al. (2021)  
generische Maskulina aktivieren die zugrundeliegenden Repräsentationen spezifischer Maskulina, was zu einer **semantischen Aktivierung spezifischer Maskulina** und damit zu einem männlichen Bias führt

